

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 26, 2005, 18:32:17 ; Search time 392.974 Seconds
(without alignments)
2797.232 Million cell updates/sec

Title: US-10-617-978-14_COPY_73_240

Perfect score: 168

Sequence: 1 ccgggaactaccacttga.....atgagaactgaaggtctga 168

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 7331713 seqs, 3271544945 residues

Word size : 0

Total number of hits satisfying chosen parameters: 14663426

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
- 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
- 13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
- 14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
- 15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
- 16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq.*
- 17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq.*
- 18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq.*
- 19: /cgn2_6/ptodata/2/pubpna/US10G_PUBCOMB.seq.*
- 20: /cgn2_6/ptodata/2/pubpna/US10H_PUBCOMB.seq.*
- 21: /cgn2_6/ptodata/2/pubpna/US10I_PUBCOMB.seq.*
- 22: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
- 23: /cgn2_6/ptodata/2/pubpna/US11A_PUBCOMB.seq.*
- 24: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
- 25: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
- 26: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	19	11.3	665	13	US-10-027-632-210235
2	19	11.3	665	17	US-10-027-632-210235
3	19	11.3	143412	13	US-10-087-192-997
4	18	10.7	195	21	US-10-721-793-139
5	18	10.7	195	21	US-10-721-793-143
6	18	10.7	195	21	US-10-721-793-147
7	18	10.7	195	21	US-10-721-793-167

8	18	10.7	195	21	US-10-721-793-171	Sequence 171, App
9	18	10.7	198	21	US-10-721-793-155	Sequence 155, App
10	18	10.7	198	21	US-10-721-793-159	Sequence 159, App
11	18	10.7	198	21	US-10-721-793-163	Sequence 163, App
12	18	10.7	272	18	US-10-424-599-1362	Sequence 13962, A
13	18	10.7	319	21	US-10-721-793-153	Sequence 153, App
14	18	10.7	323	21	US-10-721-793-137	Sequence 141, App
15	18	10.7	323	21	US-10-721-793-141	Sequence 145, App
16	18	10.7	323	21	US-10-721-793-145	Sequence 157, App
17	18	10.7	323	21	US-10-721-793-157	Sequence 161, App
18	18	10.7	323	21	US-10-721-793-161	Sequence 165, App
19	18	10.7	323	21	US-10-721-793-169	Sequence 169, App
20	18	10.7	323	21	US-10-721-793-132	Sequence 113132
21	18	10.7	3312	17	US-10-027-632-113132	Sequence 113132
22	18	10.7	3312	17	US-10-027-632-113132	Sequence 533, App
23	18	10.7	4790	17	US-10-291-172-523	Sequence 533, App
24	18	10.7	4790	17	US-10-221-278-523	Sequence 59, Appl
25	18	10.7	5587	17	US-10-085-198-59	Sequence 3132, Ap
26	18	10.7	6846	20	US-10-723-860-3132	Sequence 2924, Ap
27	18	10.7	6846	22	US-10-756-149-2924	Sequence 106, App
28	18	10.7	7037	15	US-10-119-926-106	Sequence 147, App
29	18	10.7	7037	17	US-10-291-172-147	Sequence 147, App
30	18	10.7	7037	18	US-10-221-278-147	Sequence 147, App
31	18	10.7	7132	13	US-10-044-090-780	Sequence 5, Appli
32	18	10.7	7974	19	US-10-480-172-5	Sequence 93, Appl
33	17	10.1	396	9	US-09-825-294-93	Sequence 93, Appl
34	17	10.1	396	9	US-09-970-966-93	Sequence 93, Appl
35	17	10.1	396	15	US-10-212-677-93	Sequence 93, Appl
36	17	10.1	396	17	US-10-361-811-93	Sequence 93, Appl
37	17	10.1	396	17	US-10-369-186-93	Sequence 5885, Ap
38	17	10.1	427	19	US-10-674-124A-5885	Sequence 25, Appl
39	17	10.1	468	15	US-10-338-694-25	Sequence 27823, A
40	17	10.1	489	18	US-10-242-535A-27823	Sequence 27823, A
41	17	10.1	489	18	US-10-085-783A-27823	Sequence 80648, A
42	17	10.1	600	22	US-10-972-079-80648	Sequence 256780
43	17	10.1	758	9	US-09-910-943-462	Sequence 256780
44	17	10.1	963	13	US-10-027-632-256780	Sequence 256780
45	17	10.1	963	17	US-10-027-632-256780	Sequence 256780

ALIGNMENTS

RESULT 1

US-10-027-632-210235
; Sequence 210235, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; POLYMORPHISMS IN THE HUMAN GENOME
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 210235
; LENGTH: 665
; TYPE: DNA

```
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(665)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-210235

Query Match      11.3%; Score 19; DB 13; Length 665;
Best Local Similarity 100.0%; Pred. No. 1.5;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 133 TGTGAATTTCTGAAGGATG 151
      |||||
Db 347 TGTGAATTTCTGAAGGATG 365

RESULT 2
US-10-027-632-210235
; Sequence 210235, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 10827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 210235
; LENGTH: 665
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(665)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-210235

Query Match      11.3%; Score 19; DB 17; Length 665;
Best Local Similarity 100.0%; Pred. No. 1.5;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 133 TGTGAATTTCTGAAGGATG 151
      |||||
Db 347 TGTGAATTTCTGAAGGATG 365

RESULT 3
US-10-087-192-997/c
; Sequence 997, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: CANCER
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192

; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(665)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-210235

; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 997
; LENGTH: 143412
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(143412)
; OTHER INFORMATION: n = A,T,C or G
US-10-087-192-997

Query Match      11.3%; Score 19; DB 13; Length 143412;
Best Local Similarity 100.0%; Pred. No. 1.7;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 118 GCCTTCCAATGCTGGTGTG 136
      |||||
Db 129843 GCCTTCCAATGCTGGTGTG 129825

RESULT 4
US-10-721-793-139
; Sequence 139, Application US/10721793
; Publication No. US2005006531A1
; GENERAL INFORMATION:
; APPLICANT: Corona Villegas, Miguel
; APPLICANT: Garcia Rodriguez, Ma Consuelo
; APPLICANT: Valdez Cruz, Norma Adriana
; APPLICANT: Gurrula Briones, Georgina
; APPLICANT: Becerril Lujan, Baltazar
; APPLICANT: Possani Postay, Lourival Domingos
; TITLE OF INVENTION: Recombinant Immunogens for the Generation of Antivenoms to the
; TITLE OF INVENTION: Venom of Scorpions of the Genus Centruroides
; FILE REFERENCE: 2099.0070001
; CURRENT APPLICATION NUMBER: US/10/721,793
; CURRENT FILING DATE: 2003-11-26
; PRIOR APPLICATION NUMBER: US 60/430,067
; PRIOR FILING DATE: 2002-12-02
; NUMBER OF SEQ ID NOS: 294
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 139
; LENGTH: 195
; TYPE: DNA
; ORGANISM: Centruroides sculpturatus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(195)
; OTHER INFORMATION: Product= Sodium-channel modifier toxin
; PUBLICATION INFORMATION:
; AUTHORS: Corona, M., Valdez-Cruz, N.A., Merino, E., Zurita, M. & Possani L.D.
; TITLE: Genes and peptides from the scorpion Centruroides sculpturatus Ewing,
; TITLE: that recognize Na+-channels
; JOURNAL: Toxinon
; VOLUME: 39
; ISSUE: 12
; PAGES: 1893-1898
; DATE: 2001-12-01
; DATABASE ENTRY DATE:
; RELEVANT RESIDUES: (1)...(195)
US-10-721-793-139

Query Match      10.7%; Score 18; DB 21; Length 195;
Best Local Similarity 100.0%; Pred. No. 5.5;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 102 TTACGGGTATTGCTAGC 119
      |||||
```

Db 111 TTACGGGTATTGCTACGC 128

RESULT 5

US-10-721-793-143
; Sequence 143, Application US/10721793
; Publication No. US20050065331A1
; GENERAL INFORMATION:
; APPLICANT: Corona Villegas, Miguel
; APPLICANT: Garcia Rodriguez, Ma Consuelo
; APPLICANT: Valdez Cruz, Norma Adriana
; APPLICANT: Gurrula Briones, Georgina
; APPLICANT: Becerril Lujan, Baltazar
; APPLICANT: Possani Postay, Lourival Domingos
; TITLE OF INVENTION: Recombinant Immunogens for the Generation of Antivenoms to the
; TITLE OF INVENTION: Venom of Scorpions of the Genus Centruroides
; FILE REFERENCE: 2099.0070001
; CURRENT APPLICATION NUMBER: US/10/721,793
; CURRENT FILING DATE: 2003-11-26
; PRIOR APPLICATION NUMBER: US 60/430,067
; PRIOR FILING DATE: 2002-12-02
; NUMBER OF SEQ ID NOS: 294
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 143
; LENGTH: 195
; TYPE: DNA
; ORGANISM: Centruroides sculpturatus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(195)
; OTHER INFORMATION: Product= Sodium-channel modifier toxin
; AUTHORS: Corona, M., Valdez-Cruz, N.A., Merino, E., Zurita, M. & Possani L.D.
; TITLE: Genes and peptides from the scorpion Centruroides sculpturatus Ewing,
; TITLE: That recognize Na+-channels
; JOURNAL: Toxicon
; VOLUME: 39
; ISSUE: 12
; PAGES: 1893-1898
; DATE: 2001-12-01
; DATABASE ENTRY DATE:
; RELEVANT RESIDUES: (1)..(195)
US-10-721-793-143

Query Match 10.7%; Score 18; DB 21; Length 195;
Best Local Similarity 100.0%; Pred. No. 5.5; Mismatches 0; Indels 0; Gaps 0;
Matches 18; Conservative 0;

Qy 102 TTACGGGTATTGCTACGC 119

Db 111 TTACGGGTATTGCTACGC 128

RESULT 6

US-10-721-793-147
; Sequence 147, Application US/10721793
; Publication No. US20050065331A1
; GENERAL INFORMATION:
; APPLICANT: Corona Villegas, Miguel
; APPLICANT: Garcia Rodriguez, Ma Consuelo
; APPLICANT: Valdez Cruz, Norma Adriana
; APPLICANT: Gurrula Briones, Georgina
; APPLICANT: Becerril Lujan, Baltazar
; APPLICANT: Possani Postay, Lourival Domingos
; TITLE OF INVENTION: Recombinant Immunogens for the Generation of Antivenoms to the
; TITLE OF INVENTION: Venom of Scorpions of the Genus Centruroides
; FILE REFERENCE: 2099.0070001
; CURRENT APPLICATION NUMBER: US/10/721,793
; CURRENT FILING DATE: 2003-11-26
; PRIOR APPLICATION NUMBER: US 60/430,067
; PRIOR FILING DATE: 2002-12-02
; NUMBER OF SEQ ID NOS: 294
; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 147
; LENGTH: 195
; TYPE: DNA
; ORGANISM: Centruroides sculpturatus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(195)
; OTHER INFORMATION: Product= Sodium-channel modifier toxin
; PUBLICATION INFORMATION:
; AUTHORS: Corona, M., Valdez-Cruz, N.A., Merino, E., Zurita, M. & Possani L.D.
; TITLE: Genes and peptides from the scorpion Centruroides sculpturatus Ewing,
; TITLE: That recognize Na+-channels
; JOURNAL: Toxicon
; VOLUME: 39
; ISSUE: 12
; PAGES: 1893-1898
; DATE: 2001-12-01
; DATABASE ENTRY DATE:
; RELEVANT RESIDUES: (1)..(195)
US-10-721-793-147

Query Match 10.7%; Score 18; DB 21; Length 195;
Best Local Similarity 100.0%; Pred. No. 5.5; Mismatches 0; Indels 0; Gaps 0;
Matches 18; Conservative 0;

Qy 102 TTACGGGTATTGCTACGC 119

Db 111 TTACGGGTATTGCTACGC 128

RESULT 7

US-10-721-793-167
; Sequence 167, Application US/10721793
; Publication No. US20050065331A1
; GENERAL INFORMATION:
; APPLICANT: Corona Villegas, Miguel
; APPLICANT: Garcia Rodriguez, Ma Consuelo
; APPLICANT: Valdez Cruz, Norma Adriana
; APPLICANT: Gurrula Briones, Georgina
; APPLICANT: Becerril Lujan, Baltazar
; APPLICANT: Possani Postay, Lourival Domingos
; TITLE OF INVENTION: Recombinant Immunogens for the Generation of Antivenoms to the
; TITLE OF INVENTION: Venom of Scorpions of the Genus Centruroides
; FILE REFERENCE: 2099.0070001
; CURRENT APPLICATION NUMBER: US/10/721,793
; CURRENT FILING DATE: 2003-11-26
; PRIOR APPLICATION NUMBER: US 60/430,067
; PRIOR FILING DATE: 2002-12-02
; NUMBER OF SEQ ID NOS: 294
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 167
; LENGTH: 195
; TYPE: DNA
; ORGANISM: Centruroides sculpturatus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(195)
; OTHER INFORMATION: Product= Sodium-channel modifier toxin
; PUBLICATION INFORMATION:
; AUTHORS: Corona, M., Valdez-Cruz, N.A., Merino, E., Zurita, M. & Possani L.D.
; TITLE: Genes and peptides from the scorpion Centruroides sculpturatus Ewing,
; TITLE: That recognize Na+-channels
; JOURNAL: Toxicon
; VOLUME: 39
; ISSUE: 12
; PAGES: 1893-1898
; DATE: 2001-12-01
; DATABASE ENTRY DATE:
; RELEVANT RESIDUES: (1)..(195)
US-10-721-793-167

Query Match 10.7%; Score 18; DB 21; Length 195;
Best Local Similarity 100.0%; Pred. No. 5.5; Mismatches 0; Indels 0; Gaps 0;
Matches 18; Conservative 0;

Qy 102 TTACGGGTATTGCTACGC 119

Db 111 TTACGGGTATTGCTACGC 128

Query Match 10.7%; Score 18; DB 21; Length 195;
Best Local Similarity 100.0%; Pred. No. 5.5;

```
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 102 TTACGGGTATTGCTACGC 119
   |||||
Db 111 TTACGGGTATTGCTACGC 128

RESULT 8
US-10-721-793-171
; Sequence 171, Application US/10721793
; Publication No. US20050065331A1
; GENERAL INFORMATION:
; APPLICANT: Corona Villegas, Miguel
; APPLICANT: Garcia Rodriguez, Ma Consuelo
; APPLICANT: Valdez Cruz, Norma Adriana
; APPLICANT: Gurrola Briones, Georgina
; APPLICANT: Becerril Lujan, Baltazar
; APPLICANT: Possani Postay, Lourival Domingos
; TITLE OF INVENTION: Recombinant Immunogens for the Generation of Antivenoms to the
; FILE REFERENCE: 2099.0070001
; CURRENT APPLICATION NUMBER: US/10/721.793
; PRIORITY FILING DATE: 2003-11-26
; PRIOR APPLICATION NUMBER: US 60/430,067
; PRIOR FILING DATE: 2002-12-02
; NUMBER OF SEQ ID NOS: 294
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 171
; LENGTH: 195
; TYPE: DNA
; ORGANISM: Centruroides sculpturatus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(195)
; OTHER INFORMATION: Product= Sodium-channel modifier toxin
; AUTHORS: Corona, M., Valdez-Cruz, N.A., Merino, E., Zurita, M. & Possani L.D.
; TITLE: Genes and peptides from the scorpion Centruroides sculpturatus Ewing,
; TITLE: that recognize Na+-channels
; JOURNAL: Toxicon
; VOLUME: 39
; ISSUE: 12
; PAGES: 1893-1898
; DATE: 2001-12-01
; DATABASE ENTRY DATE:
; RELEVANT RESIDUES: (1)..(195)
; QUERY MATCH 10.7%; Score 18; DB 21; Length 195;
; Best Local Similarity 100.0%; Pred. No. 5.5;
; Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 102 TTACGGGTATTGCTACGC 119
   |||||
Db 111 TTACGGGTATTGCTACGC 128

RESULT 9
US-10-721-793-155
; Sequence 155, Application US/10721793
; Publication No. US20050065331A1
; GENERAL INFORMATION:
; APPLICANT: Corona Villegas, Miguel
; APPLICANT: Garcia Rodriguez, Ma Consuelo
; APPLICANT: Valdez Cruz, Norma Adriana
; APPLICANT: Gurrola Briones, Georgina
; APPLICANT: Becerril Lujan, Baltazar
; APPLICANT: Possani Postay, Lourival Domingos
; TITLE OF INVENTION: Recombinant Immunogens for the Generation of Antivenoms to the
; FILE REFERENCE: 2099.0070001
; CURRENT APPLICATION NUMBER: US/10/721.793
; PRIORITY FILING DATE: 2003-11-26
; PRIOR APPLICATION NUMBER: US 60/430,067
; PRIOR FILING DATE: 2002-12-02
; NUMBER OF SEQ ID NOS: 294
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 159
; LENGTH: 198
; TYPE: DNA
; ORGANISM: Centruroides sculpturatus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(198)
; OTHER INFORMATION: Product= Sodium-channel modifier toxin
; AUTHORS: Corona, M., Valdez-Cruz, N.A., Merino, E., Zurita, M. & Possani L.D.
; TITLE: Genes and peptides from the scorpion Centruroides sculpturatus Ewing,
; TITLE: that recognize Na+-channels
; JOURNAL: Toxicon
; VOLUME: 39
; ISSUE: 12
; PAGES: 1893-1898
; DATE: 2001-12-01
; DATABASE ENTRY DATE:
; RELEVANT RESIDUES: (1)..(198)
; QUERY MATCH 10.7%; Score 18; DB 21; Length 195;
; Best Local Similarity 100.0%; Pred. No. 5.5;
; Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 102 TTACGGGTATTGCTACGC 119
   |||||
Db 111 TTACGGGTATTGCTACGC 128

RESULT 10
US-10-721-793-159
; Sequence 159, Application US/10721793
; Publication No. US20050065331A1
; GENERAL INFORMATION:
; APPLICANT: Corona Villegas, Miguel
; APPLICANT: Garcia Rodriguez, Ma Consuelo
; APPLICANT: Valdez Cruz, Norma Adriana
; APPLICANT: Gurrola Briones, Georgina
; APPLICANT: Becerril Lujan, Baltazar
; APPLICANT: Possani Postay, Lourival Domingos
; TITLE OF INVENTION: Recombinant Immunogens for the Generation of Antivenoms to the
; FILE REFERENCE: 2099.0070001
; CURRENT APPLICATION NUMBER: US/10/721.793
; PRIORITY FILING DATE: 2003-11-26
; PRIOR APPLICATION NUMBER: US 60/430,067
; PRIOR FILING DATE: 2002-12-02
; NUMBER OF SEQ ID NOS: 294
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 159
; LENGTH: 198
; TYPE: DNA
; ORGANISM: Centruroides sculpturatus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(198)
; OTHER INFORMATION: Product= Sodium-channel modifier toxin
; AUTHORS: Corona, M., Valdez-Cruz, N.A., Merino, E., Zurita, M. & Possani L.D.
; TITLE: Genes and peptides from the scorpion Centruroides sculpturatus Ewing,
; TITLE: that recognize Na+-channels
; JOURNAL: Toxicon
; VOLUME: 39
; ISSUE: 12
; PAGES: 1893-1898
; DATE: 2001-12-01
; DATABASE ENTRY DATE:
; RELEVANT RESIDUES: (1)..(198)
; QUERY MATCH 10.7%; Score 18; DB 21; Length 198;
; Best Local Similarity 100.0%; Pred. No. 5.5;
; Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

; PRIOR APPLICATION NUMBER: US 60/430,067
; PRIOR FILING DATE: 2002-12-02
; NUMBER OF SEQ ID NOS: 294
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 155
; LENGTH: 198
; TYPE: DNA
; ORGANISM: Centruroides sculpturatus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(198)
; OTHER INFORMATION: Product= Sodium-channel modifier toxin
; PUBLICATION INFORMATION:
; AUTHORS: Corona, M., Valdez-Cruz, N.A., Merino, E., Zurita, M. & Possani L.D.
; TITLE: Genes and peptides from the scorpion Centruroides sculpturatus Ewing,
; TITLE: that recognize Na+-channels
; JOURNAL: Toxicon
; VOLUME: 39
; ISSUE: 12
; PAGES: 1893-1898
; DATE: 2001-12-01
; DATABASE ENTRY DATE:
; RELEVANT RESIDUES: (1)..(198)
; US-10-721-793-155

Query Match 10.7%; Score 18; DB 21; Length 198;
Best Local Similarity 100.0%; Pred. No. 5.5;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 102 TTACGGGTATTGCTACGC 119
   |||||
Db 111 TTACGGGTATTGCTACGC 128

RESULT 10
US-10-721-793-159
; Sequence 159, Application US/10721793
; Publication No. US20050065331A1
; GENERAL INFORMATION:
; APPLICANT: Corona Villegas, Miguel
; APPLICANT: Garcia Rodriguez, Ma Consuelo
; APPLICANT: Valdez Cruz, Norma Adriana
; APPLICANT: Gurrola Briones, Georgina
; APPLICANT: Becerril Lujan, Baltazar
; APPLICANT: Possani Postay, Lourival Domingos
; TITLE OF INVENTION: Recombinant Immunogens for the Generation of Antivenoms to the
; FILE REFERENCE: 2099.0070001
; CURRENT APPLICATION NUMBER: US/10/721.793
; PRIORITY FILING DATE: 2003-11-26
; PRIOR APPLICATION NUMBER: US 60/430,067
; PRIOR FILING DATE: 2002-12-02
; NUMBER OF SEQ ID NOS: 294
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 159
; LENGTH: 198
; TYPE: DNA
; ORGANISM: Centruroides sculpturatus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(198)
; OTHER INFORMATION: Product= Sodium-channel modifier toxin
; PUBLICATION INFORMATION:
; AUTHORS: Corona, M., Valdez-Cruz, N.A., Merino, E., Zurita, M. & Possani L.D.
; TITLE: Genes and peptides from the scorpion Centruroides sculpturatus Ewing,
; TITLE: that recognize Na+-channels
; JOURNAL: Toxicon
; VOLUME: 39
; ISSUE: 12
; PAGES: 1893-1898
; DATE: 2001-12-01
; DATABASE ENTRY DATE:
; RELEVANT RESIDUES: (1)..(198)
; US-10-721-793-155
```


US-10-721-793-159

Query Match 10.7%; Score 18; DB 21; Length 198;
Best Local Similarity 100.0%; Pred. No. 5.5;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 102 TTACGGGTATTGCTACGC 119
|||||
DB 111 TTACGGGTATTGCTACGC 128

RESULT 11

US-10-721-793-163
; Sequence 163, Application US/10721793
; Publication No. US20050065331A1
; GENERAL INFORMATION:
; APPLICANT: Corona Villegas, Miguel
; APPLICANT: Garcia Rodriguez, Ma Consuelo
; APPLICANT: Valdez Cruz, Norma Adriana
; APPLICANT: Gurrola Briones, Georgina
; APPLICANT: Becerril Lujan, Baltazar
; APPLICANT: Possani Postay, Lourival Domingos
; TITLE OF INVENTION: Recombinant Immunogens for the Generation of Antivenoms to the
; TITLE OF INVENTION: Venom of Scorpions of the Genus Centruroides
; FILE REFERENCE: 2099.0070001
; CURRENT APPLICATION NUMBER: US/10/721,793
; CURRENT FILING DATE: 2003-11-26
; PRIOR APPLICATION NUMBER: US 60/430,067
; PRIOR FILING DATE: 2002-12-02
; NUMBER OF SEQ ID NOS: 294
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 163
; LENGTH: 198
; TYPE: DNA
; ORGANISM: Centruroides sculpturatus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(198)
; OTHER INFORMATION: Product= Sodium-channel modifier toxin
; PUBLICATION INFORMATION:
; AUTHORS: Corona, M., Valdez-Cruz, N.A., Merino, E., Zurita, M. & Possani L.D.
; TITLE: Genes and peptides from the scorpion Centruroides sculpturatus Ewing,
; JOURNAL: Toxicon
; VOLUME: 39
; ISSUE: 12
; PAGES: 1893-1898
; DATE: 2001-12-01
; DATABASE ENTRY DATE:
; RELEVANT RESIDUES: (1)..(198)
US-10-721-793-163

Query Match 10.7%; Score 18; DB 21; Length 198;
Best Local Similarity 100.0%; Pred. No. 5.5;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 102 TTACGGGTATTGCTACGC 119
|||||
DB 111 TTACGGGTATTGCTACGC 128

RESULT 12

US-10-424-599-13962/C
; Sequence 13962, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B

; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 13962
; LENGTH: 272
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_112616C.1
US-10-424-599-13962

Query Match 10.7%; Score 18; DB 18; Length 272;
Best Local Similarity 100.0%; Pred. No. 5.6;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 128 GCTGGTGTGAATTTCTGA 145
|||||
DB 96 GCTGGTGTGAATTTCTGA 79

RESULT 13

US-10-721-793-153
; Sequence 153, Application US/10721793
; Publication No. US20050065331A1
; GENERAL INFORMATION:
; APPLICANT: Corona Villegas, Miguel
; APPLICANT: Garcia Rodriguez, Ma Consuelo
; APPLICANT: Valdez Cruz, Norma Adriana
; APPLICANT: Gurrola Briones, Georgina
; APPLICANT: Becerril Lujan, Baltazar
; APPLICANT: Possani Postay, Lourival Domingos
; TITLE OF INVENTION: Recombinant Immunogens for the Generation of Antivenoms to the
; TITLE OF INVENTION: Venom of Scorpions of the Genus Centruroides
; FILE REFERENCE: 2099.0070001
; CURRENT APPLICATION NUMBER: US/10/721,793
; CURRENT FILING DATE: 2003-11-26
; PRIOR APPLICATION NUMBER: US 60/430,067
; PRIOR FILING DATE: 2002-12-02
; NUMBER OF SEQ ID NOS: 294
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 153
; LENGTH: 319
; TYPE: DNA
; ORGANISM: Centruroides sculpturatus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(261)
; OTHER INFORMATION: Product= Sodium-channel modifier toxin precursor
; OTHER INFORMATION: In the mature peptide, the last 2 basic aminoacids are cut
; FEATURE:
; NAME/KEY: 3'UTR
; LOCATION: (265)..(319)
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: (58)..(1)
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: (1)..(57)
; OTHER INFORMATION:
; PUBLICATION INFORMATION:
; AUTHORS: Corona, M., Valdez-Cruz, N.A., Merino, E., Zurita, M. & Possani L.D.
; TITLE: Genes and peptides from the scorpion Centruroides sculpturatus Ewing,
; JOURNAL: Toxicon
; VOLUME: 39
; ISSUE: 12
; PAGES: 1893-1898
; DATE: 2001-12-01
; DATABASE ENTRY DATE:
; RELEVANT RESIDUES: (1)..(261)
US-10-721-793-153

Query Match 10.7%; Score 18; DB 21; Length 319;
Best Local Similarity 100.0%; Pred. No. 5.6;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 102 TTACGGGTATTGCTACGC 119
Db 168 TTACGGGTATTGCTACGC 185
|||||

RESULT 14
US-10-721-793-137
; Sequence 137, Application US/10721793
; Publication No. US20050065331A1
; GENERAL INFORMATION:
; APPLICANT: Corona Villegas, Miguel
; APPLICANT: Garcia Rodriguez, Ma Consuelo
; APPLICANT: Valdez Cruz, Norma Adriana
; APPLICANT: Gurola Briones, Georgina
; APPLICANT: Becerril Lujan, Baltazar
; APPLICANT: Possani Postay, Lourival Domingos
; TITLE OF INVENTION: Recombinant Immunogens for the Generation of Antivenoms to the
; TITLE OF INVENTION: Venom of Scorpions of the Genus Centruroides
; FILE REFERENCE: 2099.0070001
; CURRENT APPLICATION NUMBER: US/10/721,793
; CURRENT FILING DATE: 2003-11-26
; PRIOR APPLICATION NUMBER: US 60/430,067
; PRIOR FILING DATE: 2002-12-02
; NUMBER OF SEQ ID NOS: 294
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 137
; LENGTH: 323
; TYPE: DNA
; ORGANISM: Centruroides sculpturatus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (5)..(265)
; OTHER INFORMATION: Product= Sodium-channel modifier toxin precursor
; OTHER INFORMATION: In the mature peptide, the last Cys is amidated, and the last Gly
; OTHER INFORMATION: and the last 2 basic aminoacids are cut
; FEATURE:
; NAME/KEY: 3'UTR
; LOCATION: (269)..(323)
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: 5'UTR
; LOCATION: (1)..(4)
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: (62)..()
; OTHER INFORMATION: Product= Sodium-channel modifier toxin
; OTHER INFORMATION:
; NAME/KEY: sig_peptide
; LOCATION: (5)..(61)
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: 3'UTR
; LOCATION: (269)..(323)
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: 5'UTR
; LOCATION: (1)..(4)
; OTHER INFORMATION:
; NAME/KEY: mat_peptide
; LOCATION: (62)..()
; OTHER INFORMATION: Product= Sodium-channel modifier toxin
; OTHER INFORMATION:
; NAME/KEY: sig_peptide
; LOCATION: (5)..(61)
; OTHER INFORMATION:
; OTHER INFORMATION:
; PUBLICATION INFORMATION:
; AUTHORS: Corona, M., Valdez-Cruz, N.A., Merino, E., Zurita, M. & Possani L.D.
; TITLE: Genes and peptides from the scorpion Centruroides sculpturatus Ewing,
; TITLE: that recognize Na+-channels
; JOURNAL: Toxicon
; VOLUME: 39
; ISSUE: 12
; PAGES: 1893-1898
; DATE: 2001-12-01
; DATABASE ENTRY DATE:
; RELEVANT RESIDUES: (5)..(265)
US-10-721-793-137

Query Match 10.7%; Score 18; DB 21; Length 323;
Best Local Similarity 100.0%; Pred. No. 5.6;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 102 TTACGGGTATTGCTACGC 119
Db 172 TTACGGGTATTGCTACGC 189
|||||

RESULT 15
US-10-721-793-141
; Sequence 141, Application US/10721793
; Publication No. US20050065331A1
; GENERAL INFORMATION:
; APPLICANT: Corona Villegas, Miguel
; APPLICANT: Garcia Rodriguez, Ma Consuelo
; APPLICANT: Valdez Cruz, Norma Adriana
; APPLICANT: Gurola Briones, Georgina
; APPLICANT: Becerril Lujan, Baltazar
; APPLICANT: Possani Postay, Lourival Domingos
; TITLE OF INVENTION: Recombinant Immunogens for the Generation of Antivenoms to the
; TITLE OF INVENTION: Venom of Scorpions of the Genus Centruroides
; FILE REFERENCE: 2099.0070001
; CURRENT APPLICATION NUMBER: US/10/721,793
; CURRENT FILING DATE: 2003-11-26
; PRIOR APPLICATION NUMBER: US 60/430,067
; PRIOR FILING DATE: 2002-12-02
; NUMBER OF SEQ ID NOS: 294
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 141
; LENGTH: 323
; TYPE: DNA
; ORGANISM: Centruroides sculpturatus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (5)..(265)
; OTHER INFORMATION: Product= Sodium-channel modifier toxin precursor
; OTHER INFORMATION: In the mature peptide, the last Cys is amidated, and the last Gly
; OTHER INFORMATION: and the last 2 basic aminoacids are cut
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: (62)..()
; OTHER INFORMATION: Product= Sodium-channel modifier toxin
; OTHER INFORMATION:
; NAME/KEY: sig_peptide
; LOCATION: (5)..(61)
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: 3'UTR
; LOCATION: (269)..(323)
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: 5'UTR
; LOCATION: (1)..(4)
; OTHER INFORMATION:
; OTHER INFORMATION:
; PUBLICATION INFORMATION:
; AUTHORS: Corona, M., Valdez-Cruz, N.A., Merino, E., Zurita, M. & Possani L.D.
; TITLE: Genes and peptides from the scorpion Centruroides sculpturatus Ewing,
; TITLE: that recognize Na+-channels
; JOURNAL: Toxicon
; VOLUME: 39
; ISSUE: 12
; PAGES: 1893-1898
; DATE: 2001-12-01
; DATABASE ENTRY DATE:
; RELEVANT RESIDUES: (5)..(265)
US-10-721-793-141

Query Match 10.7%; Score 18; DB 21; Length 323;
Best Local Similarity 100.0%; Pred. No. 5.6;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 102 TTACGGGTATTGCTACGC 119
Db 172 TTACGGGTATTGCTACGC 189
|||||

Search completed: August 26, 2005, 21:08:03
Job time : 395.974 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 26, 2005, 17:02:31 ; Search time 111.513 Seconds
(without alignments)
2465.133 Million cell updates/sec

Title: US-10-617-978-14_COPY_73_240

Perfect score: 168

Sequence: 1 ccgggaactaccacttga.....atgagaacgtgaaggtctga 168

Scoring table: OLIGO NUC

Gapop 60.0 , Gapext 60.0

Searched: 1202784 seqs, 818138359 residues

Word size : 0

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued Patents_NA.*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	17	10.1	396	US-09-640-173-93	Sequence 93, Appl
C 2	17	10.1	396	US-09-713-550-93	Sequence 93, Appl
C 3	17	10.1	396	US-09-825-294-93	Sequence 93, Appl
C 4	17	10.1	396	US-09-970-966-93	Sequence 93, Appl
C 5	17	10.1	1083	US-09-602-777A-259	Sequence 259, App
C 6	17	10.1	3662	US-09-907-794A-289	Sequence 289, App
C 7	17	10.1	3662	US-09-905-125A-289	Sequence 289, App
C 8	17	10.1	3662	US-09-902-775A-289	Sequence 289, App
C 9	17	10.1	3662	US-09-906-700-289	Sequence 289, App
C 10	17	10.1	3662	US-09-903-603A-289	Sequence 289, App
C 11	17	10.1	3662	US-09-904-920A-289	Sequence 289, App
C 12	17	10.1	3662	US-09-909-064-289	Sequence 289, App
C 13	17	10.1	3662	US-09-905-381A-289	Sequence 289, App
C 14	17	10.1	3662	US-09-906-618-289	Sequence 289, App
C 15	17	10.1	4053	US-09-907-794A-293	Sequence 293, App
C 16	17	10.1	4053	US-09-905-125A-293	Sequence 293, App
C 17	17	10.1	4053	US-09-902-775A-293	Sequence 293, App
C 18	17	10.1	4053	US-09-906-700-293	Sequence 293, App
C 19	17	10.1	4053	US-09-903-603A-293	Sequence 293, App
C 20	17	10.1	4053	US-09-904-920A-293	Sequence 293, App
C 21	17	10.1	4053	US-09-909-064-293	Sequence 293, App
C 22	17	10.1	4053	US-09-905-381A-293	Sequence 293, App
C 23	17	10.1	4053	US-09-906-618-293	Sequence 293, App
C 24	17	10.1	145241	US-09-949-016-17394	Sequence 17394, A
C 25	17	10.1	145241	US-09-949-016-17395	Sequence 17395, A
C 26	16	9.5	601	US-09-949-016-157499	Sequence 157499,
C 27	16	9.5	601	US-09-949-016-157606	Sequence 157606,

28	16	9.5	804	4	US-09-107-532A-2697	Sequence 2697, Ap
C 29	16	9.5	3655	3	US-08-878-474-6	Sequence 6, Appli
C 30	16	9.5	12309	4	US-09-949-016-17042	Sequence 17042, A
C 31	16	9.5	12309	4	US-09-949-016-17043	Sequence 17043, A
C 32	16	9.5	23067	4	US-09-949-016-12535	Sequence 12535, A
C 33	16	9.5	23093	4	US-09-949-016-17476	Sequence 17476, A
C 34	16	9.5	35100	1	US-08-306-691B-19	Sequence 19, Appl
C 35	16	9.5	35100	5	PCT-US93-06251-19	Sequence 19, Appl
C 36	16	9.5	40352	3	US-08-846-111D-15	Sequence 15, Appl
C 37	16	9.5	40352	3	US-09-443-077-15	Sequence 15, Appl
C 38	16	9.5	47787	4	US-09-949-016-11969	Sequence 11969, A
C 39	16	9.5	86947	4	US-09-949-016-11930	Sequence 11930, A
C 40	16	9.5	90724	4	US-09-949-016-16601	Sequence 16601, A
C 41	16	9.5	99960	4	US-09-762-311-2	Sequence 2, Appli
C 42	16	9.5	116955	4	US-09-949-016-17565	Sequence 17565, A
C 43	16	9.5	126176	4	US-09-949-016-16137	Sequence 16137, A
C 44	16	9.5	126176	4	US-09-949-016-16138	Sequence 16138, A
C 45	16	9.5	130298	4	US-09-949-016-16664	Sequence 16664, A

ALIGNMENTS

RESULT 1

US-09-640-173-93/c
; Sequence 93, Application US/09640173
; Patent No. 6613515
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Stolk, John A.
; TITLE OF INVENTION: OVARIAN TUMOR SEQUENCES AND
; TITLE OF INVENTION: METHODS OF USE THEREFOR
; FILE REFERENCE: 210121.484C2
; CURRENT APPLICATION NUMBER: US/09/640.173
; CURRENT FILING DATE: 2000-08-15
; NUMBER OF SEQ ID NOS: 196
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 93
; LENGTH: 396
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(396)
; OTHER INFORMATION: n = A,T,C or G
US-09-640-173-93

Query Match 10.1%; Score 17; DB 4; Length 396;
Best Local Similarity 100.0%; Pred. No. 9.8;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 140 TTCTGAAGGATGAGAAC 156
|||||
Db 199 TTCTGAAGGATGAGAAC 183

RESULT 2

US-09-713-550-93/c
; Sequence 93, Application US/09713550
; Patent No. 6617109
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Stolk, John A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; TITLE OF INVENTION: THERAPY AND DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210121.484C4
; CURRENT APPLICATION NUMBER: US/09/713.550
; CURRENT FILING DATE: 2000-11-14
; NUMBER OF SEQ ID NOS: 205
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 93
; LENGTH: 396
; TYPE: DNA

```

; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(396)
; OTHER INFORMATION: n = A,T,C or G
US-09-713-550-93

Query Match      10.1%; Score 17; DB 4; Length 396;
Best Local Similarity 100.0%; Pred. No. 9.8;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 140 TTCTGAAGGATGAGAAC 156
Db 199 TTCTGAAGGATGAGAAC 183

RESULT 3
US-09-825-294-93/c
; Sequence 93, Application US/09825294
; Patent No. 6710170
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Stolk, John A.
; APPLICANT: Algate, Steven P.
; TITLE OF INVENTION: Filing, Steven P.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; FILE REFERENCE: 210121.484C5
; CURRENT APPLICATION NUMBER: US/09/825,294
; CURRENT FILING DATE: 2001-04-03
; NUMBER OF SEQ ID NOS: 215
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 93
; LENGTH: 396
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(396)
; OTHER INFORMATION: n = A,T,C or G
US-09-825-294-93

Query Match      10.1%; Score 17; DB 4; Length 396;
Best Local Similarity 100.0%; Pred. No. 9.8;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 140 TTCTGAAGGATGAGAAC 156
Db 199 TTCTGAAGGATGAGAAC 183

RESULT 4
US-09-970-966-93/c
; Sequence 93, Application US/09970966
; Patent No. 6720146
; GENERAL INFORMATION:
; APPLICANT: Stolk, John A.
; APPLICANT: Molesh, David Alan
; APPLICANT: Fling, Steven P.
; APPLICANT: Xu, Jiangchun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.484C6
; CURRENT APPLICATION NUMBER: US/09/970,966
; CURRENT FILING DATE: 2001-10-02
; NUMBER OF SEQ ID NOS: 215
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 93
; LENGTH: 396
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature

```

;
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: DE 19941391.6
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: DE 19942088.2
; PRIOR FILING DATE: 1999-09-03
; NUMBER OF SEQ ID NOS: 442
; SEQ ID NO 259
; LENGTH: 1083
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (101)..(1060)
; OTHER INFORMATION: RXN01466
US-09-602-777A-259

Query Match 10.1%; Score 17; DB 4; Length 1083;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 118 GCCTTCCCAATGCTGGT 134
|||||
DB 877 GCCTTCCCAATGCTGGT 861

RESULT 6
US-09-907-794A-289/c
; Sequence 289, Application US/09907794A
; Patent No. 6635468
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnovers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kijavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/907,794A
; CURRENT FILING DATE: 2001-07-17
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090

;
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 289
; LENGTH: 3662
; TYPE: DNA
; ORGANISM: Homo Sapien
US-09-907-794A-289

Query Match 10.1%; Score 17; DB 4; Length 3662;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 140 TTCTGAGGATGAGAAC 156
|||||
DB 2818 TTCTGAGGATGAGAAC 2802

RESULT 7
US-09-905-125A-289/c
; Sequence 289, Application US/09905125A
; Patent No. 6664376
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnovers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kijavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/905,125A
; CURRENT FILING DATE: 2001-07-12
; PRIOR APPLICATION NUMBER: PCT/US00/04414

; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 289
; LENGTH: 3662
; TYPE: DNA
; ORGANISM: Homo Sapien
US-09-905-125A-289

Query Match 10.1%; Score 17; DB 4; Length 3662;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 140 TTCTGAGGATGAGAAC 156
||| ||||| ||||| |||||
Db 2818 TTCTGAGGATGAGAAC 2802

RESULT 8
US-09-902-775A-289/c
; Sequence 289, Application US/09902775A
; Patent No. 6686451
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnovers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Forq, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James

; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/902,775A
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 289
; LENGTH: 3662
; TYPE: DNA
; ORGANISM: Homo Sapien
US-09-902-775A-289

Query Match 10.1%; Score 17; DB 4; Length 3662;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 140 TTCTGAGGATGAGAAC 156
||| ||||| ||||| |||||
Db 2818 TTCTGAGGATGAGAAC 2802

RESULT 9
US-09-906-700-289/c
; Sequence 289, Application US/09906700
; Patent No. 6723535
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnovers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen

APPLICANT: Fong, Sherman
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, A.
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, Christopher J.
APPLICANT: Gurney, Austin L.
APPLICANT: Hillan, Kenneth, J.
APPLICANT: Kljavin, Ivar J.
APPLICANT: Mather, Jennie P.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William, I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: 10466-14
CURRENT APPLICATION NUMBER: US/09/906,700
PRIOR FILING DATE: 2000-09-18
PRIOR APPLICATION NUMBER: PCT/US00/04414
PRIOR FILING DATE: 2000-02-22
PRIOR APPLICATION NUMBER: US 60/143,048
PRIOR FILING DATE: 1999-07-07
PRIOR APPLICATION NUMBER: US 60/145,698
PRIOR FILING DATE: 1999-07-26
PRIOR APPLICATION NUMBER: US 60/146,222
PRIOR FILING DATE: 1999-07-28
PRIOR APPLICATION NUMBER: PCT/US99/20594
PRIOR FILING DATE: 1999-09-08
PRIOR APPLICATION NUMBER: PCT/US99/20944
PRIOR FILING DATE: 1999-09-13
PRIOR APPLICATION NUMBER: PCT/US99/21090
PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: PCT/US99/21547
PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: PCT/US99/23089
PRIOR FILING DATE: 1999-10-05
PRIOR APPLICATION NUMBER: PCT/US99/28214
PRIOR FILING DATE: 1999-11-29
PRIOR APPLICATION NUMBER: PCT/US99/28313
PRIOR FILING DATE: 1999-11-30
PRIOR APPLICATION NUMBER: PCT/US99/28564
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US99/28565
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US99/30095
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: PCT/US99/30911
PRIOR FILING DATE: 1999-12-20
PRIOR APPLICATION NUMBER: PCT/US99/30999
PRIOR FILING DATE: 1999-12-20
PRIOR APPLICATION NUMBER: PCT/US00/00219
PRIOR FILING DATE: 2000-01-05
NUMBER OF SEQ ID NOS: 423
SEQ ID NO 289
LENGTH: 3662
TYPE: DNA
ORGANISM: Homo Sapien
US-09-906-700-289

Query Match 10.1%; Score 17; DB 4; Length 3662;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 140 TTCTGAAGATGAGAAC 156
DB 2818 TTCTGAAGATGAGAAC 2802

RESULT 10
US-09-903-603A-289/c
Sequence 289 Application US/09903603A
Patent No. 6767995
GENERAL INFORMATION:
APPLICANT: Genentech, Inc.
APPLICANT: Ashkenazi, Avi
APPLICANT: Botstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Eaton, Dan L.
APPLICANT: Ferrara, Napoleone
APPLICANT: Filvaroff, Ellen
APPLICANT: Fong, Sherman
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, A.
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, Christopher J.
APPLICANT: Gurney, Austin L.
APPLICANT: Hillan, Kenneth, J.
APPLICANT: Kljavin, Ivar J.
APPLICANT: Mather, Jennie P.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William, I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: GNE.161892C12
CURRENT APPLICATION NUMBER: US/09/903,603A
CURRENT FILING DATE: 2001-07-11
PRIOR APPLICATION NUMBER: PCT/US00/04414
PRIOR FILING DATE: 2000-02-22
PRIOR APPLICATION NUMBER: US 60/143,048
PRIOR FILING DATE: 1999-07-07
PRIOR APPLICATION NUMBER: US 60/145,698
PRIOR FILING DATE: 1999-07-26
PRIOR APPLICATION NUMBER: US 60/146,222
PRIOR FILING DATE: 1999-07-28
PRIOR APPLICATION NUMBER: PCT/US99/20594
PRIOR FILING DATE: 1999-09-08
PRIOR APPLICATION NUMBER: PCT/US99/20944
PRIOR FILING DATE: 1999-09-13
PRIOR APPLICATION NUMBER: PCT/US99/21090
PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: PCT/US99/21547
PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: PCT/US99/23089
PRIOR FILING DATE: 1999-10-05
PRIOR APPLICATION NUMBER: PCT/US99/28214
PRIOR FILING DATE: 1999-11-29
PRIOR APPLICATION NUMBER: PCT/US99/28313
PRIOR FILING DATE: 1999-11-30
PRIOR APPLICATION NUMBER: PCT/US99/28564
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US99/28565
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US99/30095
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: PCT/US99/30911
PRIOR FILING DATE: 1999-12-20
PRIOR APPLICATION NUMBER: PCT/US99/30999
PRIOR FILING DATE: 1999-12-20
PRIOR APPLICATION NUMBER: PCT/US00/00219
PRIOR FILING DATE: 2000-01-05
NUMBER OF SEQ ID NOS: 423
SEQ ID NO 289
LENGTH: 3662
TYPE: DNA

; ORGANISM: Homo Sapien
US-09-603-603A-289

Query Match 10.1%; Score 17; DB 4; Length 3662;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 140 TTCTGAAGGATGAGAAC 156
|||||
Db 2818 TTCTGAAGGATGAGAAC 2802

RESULT 11

US-09-904-920A-289/c

; Sequence 289, Application US/09904920A
; Patent No. 6806352

; GENERAL INFORMATION:

; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnovers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kijavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.

; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic

; FILE REFERENCE: 10466-14

; CURRENT APPLICATION NUMBER: US/09/904,920A

; PRIOR FILING DATE: 2001-07-13

; PRIOR APPLICATION NUMBER: PCT/US00/04414

; PRIOR FILING DATE: 2000-02-22

; PRIOR APPLICATION NUMBER: US 60/143,048

; PRIOR FILING DATE: 1999-07-07

; PRIOR APPLICATION NUMBER: US 60/145,698

; PRIOR FILING DATE: 1999-07-26

; PRIOR APPLICATION NUMBER: US 60/146,222

; PRIOR FILING DATE: 1999-07-28

; PRIOR APPLICATION NUMBER: PCT/US99/20594

; PRIOR FILING DATE: 1999-09-08

; PRIOR APPLICATION NUMBER: PCT/US99/20944

; PRIOR FILING DATE: 1999-09-13

; PRIOR APPLICATION NUMBER: PCT/US99/21090

; PRIOR FILING DATE: 1999-09-15

; PRIOR APPLICATION NUMBER: PCT/US99/21547

; PRIOR FILING DATE: 1999-09-15

; PRIOR APPLICATION NUMBER: PCT/US99/23089

; PRIOR FILING DATE: 1999-10-05

; PRIOR APPLICATION NUMBER: PCT/US99/28214

; PRIOR FILING DATE: 1999-11-29

; PRIOR APPLICATION NUMBER: PCT/US99/28313

; PRIOR FILING DATE: 1999-11-30

; PRIOR APPLICATION NUMBER: PCT/US99/28564

; PRIOR FILING DATE: 1999-12-02

; PRIOR APPLICATION NUMBER: PCT/US99/28565

; PRIOR FILING DATE: 1999-12-02

; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 289
; LENGTH: 3662

; TYPE: DNA
; ORGANISM: Homo Sapien
US-09-904-920A-289

Query Match 10.1%; Score 17; DB 4; Length 3662;

Best Local Similarity 100.0%; Pred. No. 11;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 140 TTCTGAAGGATGAGAAC 156

|||||

Db 2818 TTCTGAAGGATGAGAAC 2802

RESULT 12

US-09-909-064-289/c

; Sequence 289, Application US/09909064

; Patent No. 6818449

; GENERAL INFORMATION:

; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnovers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kijavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.

; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic

; FILE REFERENCE: 10466-14

; CURRENT APPLICATION NUMBER: US/09/909,064

; PRIOR FILING DATE: 2001-07-18

; PRIOR APPLICATION NUMBER: PCT/US00/04414

; PRIOR FILING DATE: 2000-02-22

; PRIOR APPLICATION NUMBER: US 60/143,048

; PRIOR FILING DATE: 1999-07-07

; PRIOR APPLICATION NUMBER: US 60/145,698

; PRIOR FILING DATE: 1999-07-26

; PRIOR APPLICATION NUMBER: US 60/146,222

; PRIOR FILING DATE: 1999-07-28

; PRIOR APPLICATION NUMBER: PCT/US99/20594

; PRIOR FILING DATE: 1999-09-08

; PRIOR APPLICATION NUMBER: PCT/US99/20944

; PRIOR FILING DATE: 1999-09-13

; PRIOR APPLICATION NUMBER: PCT/US99/21090

; PRIOR FILING DATE: 1999-09-15

; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 289
; LENGTH: 3662
; TYPE: DNA
; ORGANISM: Homo Sapien
US-09-909-064-289

Query Match 10.1%; Score 17; DB 4; Length 3662;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 140 TTCTGAAGGATGAGAAC 156
|||||
Db 2818 TTCTGAAGGATGAGAAC 2802

RESULT 13
US-09-905-381A-289/c
; Sequence 289, Application US/09905381A
; Patent No. 6818746
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/905,381A
; CURRENT FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22

; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 289
; LENGTH: 3662
; TYPE: DNA
; ORGANISM: Homo Sapien
US-09-905-381A-289

Query Match 10.1%; Score 17; DB 4; Length 3662;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 140 TTCTGAAGGATGAGAAC 156
|||||
Db 2818 TTCTGAAGGATGAGAAC 2802

RESULT 14
US-09-906-618-289/c
; Sequence 289, Application US/09906618
; Patent No. 6828146
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.

```
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/906,618
; CURRENT FILING DATE: 2001-07-16
; PRIOR FILING DATE: 2000-02-22
; PRIOR FILING DATE: 2000-02-22
; PRIOR FILING DATE: 2000-02-22
; PRIOR FILING DATE: 1999-07-07
; PRIOR FILING DATE: 1999-07-26
; PRIOR FILING DATE: 1999-07-26
; PRIOR FILING DATE: 1999-07-28
; PRIOR FILING DATE: 1999-07-28
; PRIOR FILING DATE: 1999-07-28
; PRIOR FILING DATE: 1999-09-08
; PRIOR FILING DATE: 1999-09-13
; PRIOR FILING DATE: 1999-09-13
; PRIOR FILING DATE: 1999-09-15
; PRIOR FILING DATE: 1999-09-15
; PRIOR FILING DATE: 1999-09-15
; PRIOR FILING DATE: 1999-09-15
; PRIOR FILING DATE: 1999-10-05
; PRIOR FILING DATE: 1999-10-05
; PRIOR FILING DATE: 1999-11-29
; PRIOR FILING DATE: 1999-11-29
; PRIOR FILING DATE: 1999-11-30
; PRIOR FILING DATE: 1999-11-30
; PRIOR FILING DATE: 1999-12-02
; PRIOR FILING DATE: 1999-12-02
; PRIOR FILING DATE: 1999-12-02
; PRIOR FILING DATE: 1999-12-02
; PRIOR FILING DATE: 1999-12-16
; PRIOR FILING DATE: 1999-12-20
; PRIOR FILING DATE: 1999-12-20
; PRIOR FILING DATE: 1999-12-20
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 289
; LENGTH: 3662
; TYPE: DNA
; ORGANISM: Homo Sapien
US-09-906-618-289

Query Match 10.1%; Score 17; DB 4; Length 3662;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 140 TTCTGAAGGATGAGAAC 156
|||
Db 2818 TTCTGAAGGATGAGAAC 2802

RESULT 15
US-09-907-794A-291/c
; Sequence 293, Application US/09907794A
; Patent No. 6635468
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
```

```
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Mathew, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/907,794A
; CURRENT FILING DATE: 2001-07-17
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 293
; LENGTH: 4053
; TYPE: DNA
; ORGANISM: Homo Sapien
US-09-907-794A-293

Query Match 10.1%; Score 17; DB 4; Length 4053;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 140 TTCTGAAGGATGAGAAC 156
|||
Db 3180 TTCTGAAGGATGAGAAC 3164

Search completed: August 26, 2005, 18:46:51
```

Job time : 114.513 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 26, 2005, 18:43:08 ; Search time 308.243 Seconds
(without alignments)
3226.399 Million cell updates/sec

Title: US-10-617-978-14_COPY_73_240
Perfect score: 168
Sequence: 1 ccgggaactaccacttga.....atgagaacgtgaaggtctga 168

Scoring table: OLIGO NUC
Gapop_60.0 , Gapext 60.0

Searched: 4390206 seqs, 2959870667 residues

Word size : 0 8780412

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N_Geneseq_16Dec04.*

- 1: Geneseqn1980s.*
- 2: Geneseqn1990s.*
- 3: Geneseqn2000s.*
- 4: Geneseqn2001as.*
- 5: Geneseqn2001bs.*
- 6: Geneseqn2002as.*
- 7: Geneseqn2002bs.*
- 8: Geneseqn2003as.*
- 9: Geneseqn2003bs.*
- 10: Geneseqn2003cs.*
- 11: Geneseqn2003ds.*
- 12: Geneseqn2004as.*
- 13: Geneseqn2004bs.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	19	11.3	143412	11 ACN44512	Acn44512 Mouse gen
2	18	10.7	2365	13 ADQ85216	Adq85216 Human tum
3	18	10.7	3408	10 ADH73022	Adh73022 Human MEG
C 4	18	10.7	3408	10 ADH73024	Adh73024 Human MEG
5	18	10.7	3909	10 ADI60469	Adi60469 Secreted
6	18	10.7	4790	5 AAS45254	Aas45254 cDNA enco
7	18	10.7	5587	6 ADH48775	Adh48775 NOV25 cod
8	18	10.7	5715	10 ADH73025	Adh73025 Human MEG
9	18	10.7	6846	12 ADN04593	Adn04593 Antipsori
10	18	10.7	6846	12 ADQ20312	Adq20312 Human sof
11	18	10.7	7037	5 AAS45066	Aas45066 cDNA enco
12	18	10.7	7132	8 ABX63780	Abx63780 Human cDN
13	18	10.7	7973	10 ADD93418	Add93418 Human lip
14	18	10.7	7974	8 AAF94370	Aaf94370 Human LP2
C 15	17	10.1	396	4 AAF94902	Aaf94902 Human ova
C 16	17	10.1	396	6 ABL48852	Abi48852 Ovarian c
C 17	17	10.1	396	6 ABT03169	Abt03169 Human ova
C 18	17	10.1	396	11 ADM10762	Adm10762 Human ova
C 19	17	10.1	396	12 ADJ11092	Adj11092 Represent
C 20	17	10.1	396	12 ADM43353	Adm43353 Human ova

21	17	10.1	468	10 AADG2548	Aad62548 Human T1R
C 22	17	10.1	758	6 ABS77229	Ab977229 Frog embr
C 23	17	10.1	812	3 AAZ89353	Aaz89353 Human 18
C 24	17	10.1	1014	5 AAH67435	Aah67435 C glutami
C 25	17	10.1	1083	4 AAF71267	Aaf71267 Corynebac
26	17	10.1	1350	8 ACA01824	Aca01824 C. glutam
C 27	17	10.1	1353	5 AAH65015	Aah65015 C. glutam
C 28	17	10.1	1784	10 ADI02464	Adi02464 Human cDN
C 29	17	10.1	2500	4 AAK94785	Aak94785 Human ful
C 30	17	10.1	2500	12 ADL31859	Adl31859 Full leng
C 31	17	10.1	2667	4 ABA06424	Ab06424 Human cDN
C 32	17	10.1	2667	6 ABV83761	Abv83761 Human pol
C 33	17	10.1	2690	4 AAS28803	Aas28803 Human inn
C 34	17	10.1	2690	10 ADB31528	Adb31528 Human cDN
C 35	17	10.1	3024	12 ADQ67269	Adq67269 Novel hum
C 36	17	10.1	3598	2 AAX37725	Aax37725 Human PRO
C 37	17	10.1	3662	2 AAX52264	Aax52264 Protein P
C 38	17	10.1	3662	3 AAZ52206	Aaz52206 Human PRO
C 39	17	10.1	3662	3 ADC78601	Aad78601 Human PRO
C 40	17	10.1	3662	4 AAF72422	Aaf72422 Human PRO
C 41	17	10.1	3662	4 AAS00161	Aas00161 Human cDN
C 42	17	10.1	3662	8 ACA60239	Ac60239 Human cDN
C 43	17	10.1	3662	8 ACD07639	Ac07639 Novel hum
C 44	17	10.1	3662	8 ABX71687	Abx71687 Human cDN
C 45	17	10.1	3662	8 ACH07019	Ach07019 Human sec

ALIGNMENTS

RESULT 1

ACN44512/c
ID ACN44512 standard; DNA; 143412 BP.

AC ACN44512;

DT 18-NOV-2004 (first entry)

DE Mouse genomic sequence MCG20543.

XX Cytostatic; carcinoma; lymphoma; cancer; murine; gene; SB.

XX Mus musculus.

PN WO2003073826-A2.

PD 12-SEP-2003.

PF 28-FEB-2003; 2003WO-US006235.

PR 01-MAR-2002; 2002US-00087192.

PA (SAGR-) SAGRES DISCOVERY.

PI Morris DW;

DR WPI; 2003-328604/31.

PT Recombinant nucleic acid useful for diagnosis and treatment of carcinoma comprises a nucleotide sequence.

Claim 1; SEQ ID NO 997; Opp; English.

CC The present invention relates to novel DNA and protein sequences which are associated with carcinomas. The sequences are useful for: (i) for screening drug candidates; (ii) for screening of bioactive agent capable of binding to Carcinoma Associated Protein (CAP); (iii) for screening of a bioactive agent capable of modulating the activity of CAP; (iv) for evaluating the effect of a candidate carcinoma drug; (v) for diagnosing carcinoma; (vi) for inhibiting the activity of CAP; (vii) for treating carcinoma; (viii) for neutralizing the effect of CAP; (ix) as a biochip; (x) for diagnosing carcinoma or a propensity to carcinoma; and (xi) for determining Carcinoma Associated (CA) gene copy number. In addition, the

CC CA genes are useful as DNA vaccines and the CAP are useful as markers of
CC carcinoma including lymphoma. The present sequence is one such CA coding
CC sequence. Note: This patent is an equivalent to basic patent
CC US2002182586A1, for which no sequence data was published
XX
SQ Sequence 143412 BP; 33078 A; 32311 C; 33212 G; 34277 T; 0 U; 10534 Other;

Query Match 11.3%; Score 19; DB 11; Length 143412;
Best Local Similarity 100.0%; Pred. No. 3.4;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 118 GCCTTCCAAATGCTGCTGG 136
Db 129843 GCCTTCCAAATGCTGCTGG 129825

RESULT 2
ADQ85216
ID ADQ85216 standard; cDNA; 2365 BP.
AC ADQ85216;
XX
DT 07-OCT-2004 (first entry)
XX
DE Human tumour-associated antigenic target (TAT) cDNA sequence #2030.
XX
KW human; tumour-associated antigenic target; TAT; cytostatic; gene therapy;
KW cancer; cell proliferative disorder; gene; ss.
XX
OS Homo sapiens;
XX
PN WO2004060270-A2.
XX
PD 22-JUL-2004.
XX
PF 15-OCT-2003; 2003WO-US029126.
XX
PR 18-OCT-2002; 2002US-0418988P.
XX
PA (GETH) GENENTECH INC.
PA (WUTD/) WU T D.
PA (ZHOU/) ZHOU Y.
XX
PI Wu TD, Zhou Y;
XX
WPI; 2004-534300/51.
XX
XX
XX New nucleic acid molecule and encoded polypeptide, for diagnosing,
XX preventing or treating cell proliferative disorders such as cancer.
XX
PS Claim 1; SEQ ID NO 2030; 5504pp; English.
XX
XX The present invention describes an isolated tumour-associated antigenic
XX target (TAT) nucleic acid comprising: (a) any of 4622 nucleotide
XX sequences (see SEQ ID NO:1 to 4622); (b) the full-length coding region of
XX (a); (c) the complement of (a) or (b); (d) a sequence that has 80%
XX sequence identity to (a)-(c); or (e) a sequence that hybridises to (a)-
XX nucleic acid; (2) a host cell comprising the above expression vector; (3)
XX a process for producing a polypeptide; (4) an isolated polypeptide
XX comprising: (a) an amino acid sequence encoded by any of the above
XX nucleotide sequences; (b) an amino acid sequence encoded by the full-
XX length coding region of the above nucleotide sequences; or (c) a sequence
XX having at least 80% identical to (a) or (b); (5) a chimeric polypeptide
XX comprising the above polypeptide fused to a heterologous polypeptide; (6)
XX an isolated antibody that binds to the above polypeptide; (7) a process
XX for producing the antibody; (8) an isolated oligopeptide that binds to
XX the above polypeptide; (9) a tumour-associated antigenic target (TAT)
XX binding organic molecule that binds to the above polypeptide; (10) a
XX composition of matter comprising the above (chimeric) polypeptide,
XX antibody, oligopeptide or TAT binding organic molecule, in combination
XX with a carrier; (11) an article of manufacture comprising a container and
XX the composition of matter contained within the container; (12) methods of

CC inhibiting the growth of a cell that expresses the above protein, where
CC the growth of the cell is at least in part dependent upon a growth
CC potentiating effect of the above protein; (13) a method of
CC therapeutically treating a mammal having a cancerous tumour comprising
CC cells that express the above protein; (14) a method of determining the
CC presence of a protein in a sample suspected of containing the protein
CC described above; (15) methods of diagnosing the presence of a tumour in a
CC mammal; (16) a method for treating or preventing a cell proliferative
CC disorder associated with increased expression or activity of the above
CC protein; and (17) a method of binding an antibody, oligopeptide or
CC organic molecule to a cell that expresses the protein described above.
CC The TAT sequences have cytostatic activities, and can be used in gene
CC therapy. The composition and methods are useful for diagnosing,
CC preventing or treating cancer. The composition is also used for preparing
CC a medicament for the therapeutic treatment or diagnostic detection of a
CC cell proliferative disorder or cancer. The present sequence represents a
CC human TAT cDNA sequence from the present invention.
XX
SQ Sequence 2365 BP; 554 A; 656 C; 684 G; 471 T; 0 U; 0 Other;

Query Match 10.7%; Score 18; DB 13; Length 2365;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 121 TTCCAAATGCTGCTGGTGA 138
Db 367 TTCCAAATGCTGCTGGTGA 384

RESULT 3
ADH73022
ID ADH73022 standard; cDNA; 3408 BP.
XX
AC ADH73022;
XX
DT 25-MAR-2004 (first entry)
XX
DE Human MEGF7-related cDNA sequence SeqId1.
XX
KW MEGF7; epidermal growth factor-like domain; low density lipoprotein;
KW LDL receptor-like protein; gene therapy; protein therapy;
KW MEGF7 expression; gene; ss; human.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..3408
FT /tag= a
FT /product= "Human MEGF7-related protein SeqId2"
FT /partial
FT /note= "No stop codon"
XX
XX GB2381790-A.
XX
PD 14-MAY-2003.
XX
PF 26-SEP-2002; 2002GB-00022372.
XX
PR 26-SEP-2001; 2001GB-00023124.
PR 26-JUN-2002; 2002GB-00014703.
XX
PA (GLAX) GLAXO GROUP LTD.
XX
XX Volpe F;
XX
XX WPI; 2003-432835/41.
XX P-PSDB; ADH73023.
XX
XX Novel isolated human MEGF7 polypeptide, a polypeptide having multiple
XX epidermal growth factor-like domains, useful in diagnostic assays for
XX detecting diseases associated with inappropriate MEGF7 activity or
XX levels.

PS Disclosure; SEQ ID NO 1; 47pp; English.

XX This invention relates to a novel human MEGF7 polypeptide (a polypeptide having multiple epidermal growth factor-like domains, including an unidentifiable low density lipoprotein (LDL) receptor-like protein). The sequences of the invention may be useful for gene therapy or protein therapy or for the modulation of MEGF7 expression and activity. The invention may be useful for the development of methods to diagnose or treat diseases associated with inappropriate MEGF7 activity or levels. The polypeptides and polynucleotides are also useful for configuring screening methods for detecting the effect of added compounds on the production of mRNA and polypeptide in cells. The polynucleotides are also useful as diagnostic reagents, through detecting mutations in the associated gene. The polynucleotide sequence is useful for chromosome localisation studies and tissue expression studies. The present sequence is that of a cDNA sequence which is related to the invention.

XX Sequence 3408 BP; 752 A; 943 C; 1028 G; 685 T; 0 U; 0 Other;

Query Match 10.7%; Score 18; DB 10; Length 3408;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 121 TTCCAATGCTGCTGTGAA 138
|||||

Db 1411 TTCCAATGCTGCTGTGAA 1428
|||||

RESULT 4

ADH73024/c

ID ADH73024 standard; DNA; 3408 BP.

XX ADH73024;

XX 25-WAR-2004 (first entry)

DE Human MEGF7-related DNA sequence SeqID3.

XX MEGF7; epidermal growth factor-like domain; low density lipoprotein; LDL receptor-like protein; gene therapy; protein therapy;

KW MEGF7 expression; ds; human.

XX Homo sapiens.

XX GB2381790-A.

XX 14-MAY-2003.

XX 26-SEP-2002; 2002GB-00022372.

XX 26-SEP-2001; 2001GB-00023124.

PR 26-JUN-2002; 2002GB-00014703.

XX (GLAX) GLAXO GROUP LTD.

XX Volpe F;

XX WPI; 2003-432835/41.

XX Novel isolated human MEGF7 polypeptide, a polypeptide having multiple epidermal growth factor-like domains, useful in diagnostic assays for detecting diseases associated with inappropriate MEGF7 activity or levels.

PS Disclosure; SEQ ID NO 3; 47pp; English.

XX This invention relates to a novel human MEGF7 polypeptide (a polypeptide having multiple epidermal growth factor-like domains, including an unidentifiable low density lipoprotein (LDL) receptor-like protein). The sequences of the invention may be useful for gene therapy or protein therapy or for the modulation of MEGF7 expression and activity. The invention may be useful for the development of methods to diagnose or treat diseases associated with inappropriate MEGF7 activity or levels.

CC The polypeptides and polynucleotides are also useful for configuring screening methods for detecting the effect of added compounds on the production of mRNA and polypeptide in cells. The polynucleotides are also useful as diagnostic reagents, through detecting mutations in the associated gene. The polynucleotide sequence is useful for chromosome localisation studies and tissue expression studies. The present sequence is that of a DNA sequence which is related to the invention.

SQ Sequence 3408 BP; 685 A; 1028 C; 943 G; 752 T; 0 U; 0 Other;

Query Match 10.7%; Score 18; DB 10; Length 3408;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 121 TTCCAATGCTGCTGTGAA 138
|||||

Db 1998 TTCCAATGCTGCTGTGAA 1981
|||||

RESULT 5

ADI60469

ID ADI60469 standard; DNA; 3909 BP.

XX ADI60469;

XX 15-APR-2004 (first entry)

DE Secreted polypeptide encoding gene #8.

XX ds; gene; osteopathic; vulnery; cytostatic; gene therapy; diagnosis; forensics; gene mapping; mutation identification; biodiversity;

KW chromosome marker; immune response; myeloid cell disorder;

KW lymphoid cell disorder; bone cartilage; tendon; ligament;

KW nerve tissue growth; wound healing; burns; incision; ulcer; cancer.

XX Homo sapiens.

XX WO2003025142-A2.

XX 27-MAR-2003.

XX 18-SEP-2002; 2002WO-US029636.

PR 18-SEP-2001; 2001US-0323349P.

PR 16-SEP-2002; 2002US-00323349.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Asundi V, Goodrich RW, Ren F, Zhang J, Zhao QA, Wang J; Ghosh M, Xue AJ, Wehrman T, Weng G, Zhou P, Drmanac RT;

XX WPI; 2003-354601/33.

DR P-PSDB; ADI60124.

XX New polynucleotides and secreted proteins, useful for treating myeloid or lymphoid cell disorders, in bone cartilage, tendon, ligament and nerve tissue growth or regeneration, in wound healing, and in tissue repair and replacement.

PT Claim 1; SEQ ID NO 8; 243pp; English.

PS The invention relates to novel isolated polynucleotides or a sequence encoding a polypeptide with biological activity, where the polynucleotide hybridizes to the polynucleotide under stringent hybridization conditions or has greater than 99% sequence identity with the polynucleotide. The polynucleotides and polypeptides are useful in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders and other traits, to assess biodiversity, as nutritional sources or supplements. The polynucleotides may also be used as molecular weight markers, chromosome markers or map related gene positions, or as an antigen to raise anti-DNA antibodies or elicit immune response. The polypeptides are useful for raising antibodies, as markers for tissues in which the corresponding polypeptide is expressed, for re-engineering

CC damaged or diseased tissues, for treating myeloid or lymphoid cell
CC disorders, in bone cartilage, tendon, ligament and/or nerve tissue growth
CC or regeneration, in wound healing, in tissue repair and replacement, in
CC healing of burns, incisions and ulcers, and in treating cancer. This
CC sequence corresponds to a polynucleotide sequence of the invention.
XX
SQ Sequence 3909 BP; 826 A; 1092 C; 1168 G; 823 T; 0 U; 0 Other;

Query Match 10.7%; Score 18; DB 10; Length 3909;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 121 TTCCAATGCTGGTGTGAA 138
|||||
Db 1507 TTCCAATGCTGGTGTGAA 1524

RESULT 6
AAS45254
ID AAS45254 standard; cDNA; 4790 BP.

AC AAS45254;

XX 18-DEC-2001 (first entry)

XX cDNA encoding novel human secretory protein, Seq ID No 523.

XX Human; secreted protein; arthritis; Crohn's disease; sepsis; shock;
KW ischaemia-reperfusion injury; haematopoiesis; cancer; neuropathy;
KW transgenic animal; Alzheimer's disease; Parkinson's disease; burn;
KW amyotrophic lateral sclerosis; platelet disorder; thrombocytopenia;
KW ulcer; osteoporosis; bone degenerative disorder; periodontal disease;
KW gut protection; lung; liver fibrosis; immune deficiency; infection;
KW severe combined immunodeficiency; SCID; autoimmune disorder; allergy;
KW multiple sclerosis; rheumatoid arthritis; diabetes mellitus; asthma;
KW fertility; analgesic; pain; antigen; ss.

XX Homo sapiens.

XX WO200166689-A2.

XX 13-SEP-2001.

XX 05-MAR-2001; 2001WO-US004942.

XX 07-MAR-2000; 2000US-00519705.

PR 19-MAY-2000; 2000US-00574454.

PR 17-JUN-2000; 2000US-00596193.

PR 14-JUL-2000; 2000US-00616847.

PR 19-SEP-2000; 2000US-00665363.

PR 20-OCT-2000; 2000US-00693267.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Asundi V, Xu C, Wehrman T, Ren F, Ma Y, Zhou P;

PI Zhao QA, Yang Y, Drmanac RT, Zhang J, Chen R, Xue AJ, Wang J;

XX WPI; 2001-589934/66.

DR P-PSDB; AAU28354.

XX Novel polypeptides and nucleic acids obtained from cDNA libraries

PT prepared from various human tissues, for diagnosis and treatment of

PT cancer, neurological, inflammatory, and autoimmune disorders.

XX Claim 1; SEQ ID NO 523; 107pp; English.

XX The invention relates to novel isolated human secreted polypeptides (I)

CC and polynucleotides (II). (I) and (II) are useful for treating

CC inflammatory conditions such as arthritis, nephritis, Crohn's disease,

CC ischaemia-reperfusion injury, shock, sepsis, immune responses, and is

CC involved in increasing haematopoiesis, stem cell survival, bone growth

CC and remodeling. (I), (II) and modulators of (II) are useful for

CC prophylaxis or treatment of one or more cancers. (II) is also useful for

CC creating transgenic animals useful for studying the in vivo activities of
CC the polypeptide as well as for studying modulators of the polypeptides.
CC (I) induces the proliferation of neural cells and regeneration of nerve
CC and brain tissue and is useful for the treatment of central and
CC peripheral nervous system diseases and neuropathies, such as Alzheimer's,
CC Parkinson's disease, Huntington's disease, and amyotrophic lateral
CC sclerosis. In addition, (I) is involved in chemotactic or chemokinetic
CC activity, regulation of haematopoiesis and is useful for treating myeloid
CC or lymphoid cell disorders, platelet disorders such as thrombocytopenia
CC and for regeneration of bone, cartilage, tendon, ligament and/or nerve
CC tissue growth, and in tissue repair, healing of burns, incisions, ulcers,
CC for treating osteoporosis, osteoarthritis, bone degenerative disorders,
CC or periodontal disease. Furthermore, (I) is also useful for gut
CC protection or regeneration and treatment of lung or liver fibrosis, and
CC reperfusion injury in various tissues, various immune deficiencies and
CC disorders including severe combined immunodeficiency (SCID), bacterial or
CC fungal infections, autoimmune disorders e.g. multiple sclerosis,
CC rheumatoid arthritis, diabetes mellitus, myasthenia gravis, allergic
CC reactions and conditions, such as asthma or other respiratory problems.
CC In addition, (I) affects biorhythms or circadian cycles of rhythms, of
CC fertility, metabolism, catabolism, anabolism, storage or elimination of
CC dietary fat, lipid, protein, carbohydrate, vitamins, minerals, provides
CC analgesic effects or other pain reducing effects, immunoglobulin like
CC activity and can act as an antigen in a vaccine composition to raise an
CC immune response. AAS44920-AAS45295 represent novel human secreted protein
CC coding sequences of the invention
XX

SQ Sequence 4790 BP; 1126 A; 1336 C; 1368 G; 960 T; 0 U; 0 Other;

Query Match 10.7%; Score 18; DB 5; Length 4790;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 121 TTCCAATGCTGGTGTGAA 138
|||||
Db 321 TTCCAATGCTGGTGTGAA 338

RESULT 7

ADH48775

ID ADH48775 standard; DNA; 5587 BP.

AC ADH48775;

XX 25-MAR-2004 (first entry)

DT NOV25 coding sequence, SEQ ID 59.

XX Human; NOVX; atherosclerosis; hypertension; obesity; cancer; cytostatic;

KW hypotensive; antiarteriosclerotic; anorectic; gene therapy; NOV25;

KW MGF7-like protein; chromosome 11; gene; ds; SNP;

KW single nucleotide polymorphism.

XX Homo sapiens.

XX Key Location/Qualifiers

FT Variation replace(1836,T)

FT /*tag= a

FT /standard_name= "Single nucleotide polymorphism"

XX WO200268652-A2.

XX 06-SEP-2002.

XX 26-FEB-2002; 2002WO-US005910.

XX 26-FEB-2001; 2001US-0271646P.

PR 28-FEB-2001; 2001US-0271840P.

PR 28-FEB-2001; 2001US-0272404P.

PR 28-FEB-2001; 2001US-0272405P.

PR 28-FEB-2001; 2001US-0272410P.

PR 28-FEB-2001; 2001US-0272414P.

PR 02-MAR-2001; 2001US-0272787P.

```

PR 02-MAR-2001; 2001US-0272922P.
PR 02-MAR-2001; 2001US-0273048P.
PR 02-MAR-2001; 2001US-0273300P.
PR 16-MAR-2001; 2001US-0276401P.
PR 20-MAR-2001; 2001US-0277324P.
PR 20-MAR-2001; 2001US-0278660P.
PR 30-MAR-2001; 2001US-0280039P.
PR 30-MAR-2001; 2001US-0280234P.
PR 02-APR-2001; 2001US-0280818P.
PR 12-APR-2001; 2001US-0283443P.
PR 23-APR-2001; 2001US-0285754P.
PR 24-APR-2001; 2001US-0286098P.
PR 03-MAY-2001; 2001US-0288353P.
PR 17-MAY-2001; 2001US-0291703P.
PR 31-MAY-2001; 2001US-0294834P.
PR 20-JUN-2001; 2001US-0299695P.
PR 21-JUN-2001; 2001US-0299845P.
PR 05-JUL-2001; 2001US-030242P.
PR 13-AUG-2001; 2001US-0311981P.
PR 16-AUG-2001; 2001US-0312858P.
PR 17-AUG-2001; 2001US-0313280P.
PR 29-AUG-2001; 2001US-0315614P.
PR 17-SEP-2001; 2001US-0322818P.
PR 25-FEB-2002; 2002US-00322818.
XX
PA (CURA-) CURAGEN CORP.
XX
XX Alsbrook JP, Anderson DW, Ballinger RA, Boldog FL, Burgess CE;
PI Casman SJ, Ellerman KE, Gangolli EA, Gerlach VL, Gilbert JA;
PI Gorman L, Guo X, Gusev VI, Kekuda R, Li L, Liu X, Malyankar UM;
PI Miller CE, Millet I, Padigaru M, Patturajan M, Pena CEA, Peyman JA;
PI Rastelli L, Shenoy SG, Shimkets RA, Smithson G, Spytek KA, Stone DJ;
PI Taupier RJ, Tchernev VT, Vernet CAM, Zerhusen BD;
XX
XX WPI; 2002-698672/75.
DR P-PSDB; ADH48776.
XX
XX New NOVX polypeptides or polynucleotides, useful for preventing or
PT treating disorders or syndromes e.g., atherosclerosis, hypertension,
PT obesity or cancer.
XX
XX Claim 8; Page 156-157; 923pp; English.
PS
XX The present invention relates to novel human NOVX proteins, where X is
CC any number from 1 to 91 and their coding sequences. The proteins and
CC coding sequences are useful for preventing or treating disorders or
CC syndromes e.g. atherosclerosis, hypertension, obesity or cancer. NOV25 is
CC a MEGF7-like protein and its coding sequence maps to chromosome 11.
XX
XX Sequence 5587 BP; 1297 A; 1556 C; 1610 G; 1124 T; 0 U; 0 Other;
SQ
Query Match 10.7%; Score 18; DB 6; Length 5587;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 121 TTCCAATGCTGGTGTGAA 138
DB 1087 TTCCAATGCTGGTGTGAA 1104
RESULT 8
ADH73025
ID ADH73025 standard; DNA; 5715 BP.
XX
XX AC ADH73025;
XX
XX 25-MAR-2004 (first entry)
XX
XX Human MEGF7 gene sequence.
XX
XX MEGF7; epidermal growth factor-like domain; low density lipoprotein;
KW LDL receptor-like protein; gene therapy; protein therapy;
KW MEGF7 expression; human; ds; gene.
XX

```

```

XX OS Homo sapiens.
XX
XX FH Key Location/Qualifiers
XX CDS 1..5715
XX FT /*tag= a
XX FT /product= "Human MEGF7 protein"
XX FT /partial
XX FT /note= "No stop codon"
XX
XX GB2381790-A.
XX
XX PD 14-MAY-2003.
XX
XX 26-SEP-2002; 2002GB-00022372.
XX
XX 26-SEP-2001; 2001GB-00023124.
XX
XX 26-JUN-2002; 2002GB-00014703.
XX
XX (GLAX ) GLAXO GROUP LTD.
XX
XX Volpe F;
XX
XX WPI; 2003-432835/41.
DR P-PSDB; ADH73026.
XX
XX Novel isolated human MEGF7 polypeptide, a polypeptide having multiple
PT epidermal growth factor-like domains, useful in diagnostic assays for
PT detecting diseases associated with inappropriate MEGF7 activity or
PT levels.
XX
XX Claim 1; SEQ ID NO 4; 47pp; English.
PS
XX This invention relates to a novel human MEGF7 polypeptide (a polypeptide
CC having multiple epidermal growth factor-like domains, including an
CC unidentified low density lipoprotein (LDL) receptor-like protein). The
CC sequences of the invention may be useful for gene therapy or protein
CC therapy or for the modulation of MEGF7 expression and activity. The
CC invention may be useful for the development of methods to diagnose or
CC treat diseases associated with inappropriate MEGF7 activity or levels.
CC The polypeptides and polynucleotides are also useful for configuring
CC screening methods for detecting the effect of added compounds on the
CC production of mRNA and polypeptide in cells. The polynucleotides are also
CC useful as diagnostic reagents, through detecting mutations in the
CC associated gene. The polynucleotide sequence is useful for chromosome
CC localisation studies and tissue expression studies. The present sequence
CC is the gene which encodes the human MEGF7 protein of the invention.
XX
XX Sequence 5715 BP; 1316 A; 1588 C; 1665 G; 1146 T; 0 U; 0 Other;
SQ
Query Match 10.7%; Score 18; DB 10; Length 5715;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 121 TTCCAATGCTGGTGTGAA 138
DB 1246 TTCCAATGCTGGTGTGAA 1263
RESULT 9
ADN04593
ID ADN04593 standard; cDNA; 6846 BP.
XX
XX AC ADN04593;
XX
XX 01-JUL-2004 (first entry)
XX
XX Antipsoriatic cDNA sequence #504.
XX
XX ds; gene; antipsoriatic; gene therapy; psoriasis; diagnosis.
KW
XX Homo sapiens.
XX

```

```

PN WO2004028479-A2.
PD
XX 08-APR-2004.
XX
XX 25-SEP-2003; 2003WO-US030907.
XX
XX 25-SEP-2002; 2002US-0414006P.
XX
XX (GETH ) GENENTECH INC.
XX
XX Bodary S, Clark H, Jackman J, Schoenfeld J, Williams PM, Wood WI;
XX Wu TD;
XX
XX WPI; 2004-305105/28.
XX
XX P-PSDB; ADN04594.
XX
XX New PRO nucleic acid or polypeptide, useful for preparing a
XX pharmaceutical composition for diagnosing or treating psoriasis in a
XX mammal.
XX
XX Claim 1; SEQ ID NO 987; 3069pp; English.
XX
XX The invention relates to novel polynucleotide and polypeptides for
XX treating psoriasis or a sequence having at least 80% identity to the
XX above sequences. The nucleic acid is useful for preparing a composition
XX for diagnosing or treating psoriasis in a mammal. This sequence
XX corresponds to one of the polynucleotides of the invention.
XX
XX SQ Sequence 6846 BP; 1640 A; 1818 C; 1862 G; 1526 T; 0 U; 0 Other;
XX
XX Query Match 10.7%; Score 18; DB 12; Length 6846;
XX Best Local Similarity 100.0%; Pred. No. 12;
XX Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX Qy 121 TTCCAATGCTGGTGTGAA 138
XX |||||||||||||||
XX Db 259 TTCCAATGCTGGTGTGAA 276
XX
XX RESULT 11
XX AAS45066
XX ID AAS45066 standard; cDNA; 7037 BP.
XX
XX AC AAS45066;
XX
XX DT 18-DEC-2001 (first entry)
XX
XX DE cDNA encoding novel human secretory protein, Seq ID No 147.
XX
XX KW Human; secreted protein; arthritis; Crohn's disease; sepsis; shock;
XX ischaemia-reperfusion injury; haematopoiesis; cancer; neuropathy;
XX transgenic animal; Alzheimer's disease; Parkinson's disease; burn;
XX amyotrophic lateral sclerosis; platelet disorder; thrombocytopenia;
XX ulcer; osteoporosis; bone degenerative disorder; periodontal disease;
XX gut protection; lung; liver fibrosis; immune deficiency; infection;
XX severe combined immunodeficiency; SCID; autoimmune disorder; allergy;
XX multiple sclerosis; rheumatoid arthritis; diabetes mellitus; asthma;
XX fertility; analgesic; pain; antigen; ss.
XX
XX OS Homo sapiens.
XX
XX PN WO200166689-A2.
XX
XX PD 13-SEP-2001.
XX
XX PF 05-MAR-2001; 2001WO-US004942.
XX
XX PR 07-MAR-2000; 2000US-00519705.
XX
XX PR 19-MAY-2000; 2000US-00574454.
XX
XX PR 17-JUN-2000; 2000US-00596193.
XX
XX PR 14-JUL-2000; 2000US-00616847.
XX
XX PR 19-SEP-2000; 2000US-00665363.
XX
XX PR 20-OCT-2000; 2000US-00693267.
XX
XX (HYSE-) HYSEQ INC.
XX
XX PA Tang YT, Liu C, Asundi V, Xu C, Wehrman T, Ren F, Ma Y, Zhou P;
XX Zhao QA, Yang Y, Drmanac RT, Zhang J, Chen R, Xue AJ, Wang J;
XX
XX WPI; 2001-589934/66.
XX
XX DR P-PSDB; AAU28166.
XX
XX PT Novel polypeptides and nucleic acids obtained from cDNA libraries
XX prepared from various human tissues, for diagnosis and treatment of
XX cancer, neurological, inflammatory, and autoimmune disorders.
XX

```

```

PN WO2004028479-A2.
PD
XX 08-APR-2004.
XX
XX 25-SEP-2003; 2003WO-US030907.
XX
XX 25-SEP-2002; 2002US-0414006P.
XX
XX (GETH ) GENENTECH INC.
XX
XX Bodary S, Clark H, Jackman J, Schoenfeld J, Williams PM, Wood WI;
XX Wu TD;
XX
XX WPI; 2004-305105/28.
XX
XX P-PSDB; ADN04594.
XX
XX New PRO nucleic acid or polypeptide, useful for preparing a
XX pharmaceutical composition for diagnosing or treating psoriasis in a
XX mammal.
XX
XX Claim 1; SEQ ID NO 987; 3069pp; English.
XX
XX The invention relates to novel polynucleotide and polypeptides for
XX treating psoriasis or a sequence having at least 80% identity to the
XX above sequences. The nucleic acid is useful for preparing a composition
XX for diagnosing or treating psoriasis in a mammal. This sequence
XX corresponds to one of the polynucleotides of the invention.
XX
XX SQ Sequence 6846 BP; 1640 A; 1818 C; 1862 G; 1526 T; 0 U; 0 Other;
XX
XX Query Match 10.7%; Score 18; DB 12; Length 6846;
XX Best Local Similarity 100.0%; Pred. No. 12;
XX Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX Qy 121 TTCCAATGCTGGTGTGAA 138
XX |||||||||||||||
XX Db 259 TTCCAATGCTGGTGTGAA 276
XX
XX RESULT 10
XX ADQ20312
XX ID ADQ20312 standard; DNA; 6846 BP.
XX
XX AC ADQ20312;
XX
XX DT 26-AUG-2004 (first entry)
XX
XX DE Human soft tissue sarcoma-upregulated DNA - SEQ ID 3132.
XX
XX KW soft tissue sarcoma; cytostatic; gene therapy; vaccine; screening; human;
XX ds.
XX
XX OS Homo sapiens;
XX
XX PN WO2004048938-A2.
XX
XX PD 10-JUN-2004.
XX
XX PF 26-NOV-2003; 2003WO-US038193.
XX
XX PR 26-NOV-2002; 2002US-0429739P.
XX
XX (PROT-) PROTEIN DESIGN LABS INC.
XX
XX PA Aziz N, Ginsburg WM, Zlotnik A;
XX
XX WPI; 2004-441208/41.
XX
XX PT Early detection of soft tissue sarcoma comprises determining expression
XX of a gene in a first soft tissue sample and a normal soft tissue sample
XX and comparing the gene expression, also useful in treating soft tissue
XX sarcoma.
XX

```

XX Claim 1; SEQ ID NO 147; 107pp; English.

XX The invention relates to novel isolated human secreted polypeptides (I)

XX and polynucleotides (II). (I) and (II) are useful for treating

CC inflammatory conditions such as arthritis, nephritis, Crohn's disease,

CC ischaemia-reperfusion injury, shock, sepsis, immune responses, and is

CC involved in increasing haematopoiesis, stem cell survival, bone growth

CC and remodeling. (I), (II) and modulators of (II) are useful for

CC prophylaxis or treatment of one or more cancers. (II) is also useful for

CC creating transgenic animals useful for studying the in vivo activities of

CC the polypeptide as well as for studying modulators of the polypeptides.

CC (I) induces the proliferation of neural cells and regeneration of nerve

CC and brain tissue and is useful for the treatment of central and

CC peripheral nervous system diseases and neuropathies, such as Alzheimer's,

CC Parkinson's disease, Huntington's disease, and amyotrophic lateral

CC sclerosis. In addition, (I) is involved in chemotactic or chemokinetic

CC activity, regulation of haematopoiesis and is useful for treating myeloid

CC or lymphoid cell disorders, platelet disorders such as thrombocytopenia

CC and for regeneration of bone, cartilage, tendon, ligament and/or nerve

CC tissue growth, and in tissue repair, healing of burns, incisions, ulcers,

CC for treating osteoporosis, osteoarthritis, bone degenerative disorders,

CC or periodontal disease. Furthermore, (I) is also useful for gut

CC protection or regeneration and treatment of lung or liver fibrosis,

CC reperfusion injury in various tissues, various immune deficiencies and

CC disorders including severe combined immunodeficiency (SCID), bacterial or

CC fungal infections, autoimmune disorders e.g. multiple sclerosis,

CC rheumatoid arthritis, diabetes mellitus, myasthenia gravis, allergic

CC reactions and conditions, such as asthma or other respiratory problems.

CC In addition, (I) affects biorhythms or circadian cycles of rhythms,

CC fertility, metabolism, catabolism, anabolism, storage or elimination of

CC dietary fat, lipid, protein, carbohydrate, vitamins, minerals, provides

CC analgesic effects or other pain reducing effects, immunoglobulin like

CC activity and can act as an antigen in a vaccine composition to raise an

CC immune response. AAS44920-AAS45295 represent novel human secreted protein

CC coding sequences of the invention

XX

XX Sequence 7037 BP; 1691 A; 1867 C; 1925 G; 1564 T; 0 U; 0 Other;

XX

Query Match 10.7%; Score 18; DB 5; Length 7037;

Best Local Similarity 100.0%; Pred. No. 12;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX

QY 121 TTCCAATGCTGGTGTGAA 138

Db 450 TTCCAATGCTGGTGTGAA 467

XX

RESULT 12

ID ABX63780

XX ABX63780 standard; cDNA; 7132 BP.

XX

AC ABX63780;

XX

XX 26-FEB-2003 (first entry)

XX

XX Human cDNA #780 differentially expressed in activated vascular tissue.

XX

XX Human; gene; ss; vascular tissue; cytostatic; atherosclerosis; cardiant;

XX hypotensive; antidiabetic; gynaecological; vasotropic; cerebroprotective;

XX gene therapy; vascular disease; cancer; coronary; artery disease;

XX hypertension; diabetes; pre-eclampsia; restenosis;

XX ischaemia-reperfusion injury; stroke.

XX

XX Homo sapiens.

OS

XX US2002137081-A1.

XX

XX 26-SEP-2002.

PD

XX 08-JAN-2002; 2002US-00044090.

PF

XX 28-JUL-2000; 2000US-0222469P.

PR

PR 08-JAN-2001; 2001US-0260483P.

XX

XX (BAND/) BANDMAN O.

PI

XX Bandman O;

XX

DR WPI; 2003-110597/10.

XX

PT Combination for diagnosing, staging, treating, or monitoring the

PT progression of treatment of a vascular disease, e.g. atherosclerosis,

PT comprises several cDNAs that are differentially expressed in activated

PT vascular tissue.

XX

PS Claim 1; Page; 18pp; English.

XX

XX This invention relates to a combination comprising several cDNAs that are

XX differentially expressed in activated vascular tissue. The invention also

XX discloses a high throughput method for detecting differentially expressed

XX cDNAs in a sample. The cDNAs of the invention may have

CC antiarteriosclerotic; cytostatic; cardiant; hypotensive; antidiabetic;

CC gynaecological; vasotropic and cerebroprotective activities and may be

CC used in gene therapy. The cDNAs of the invention may be used in a high-

CC throughput methods for detecting differential expression of one or more

CC cDNAs in a sample, or screening several molecules or compounds to

CC identify a molecule or compound that specifically binds a cDNA of the

CC invention. A protein encoded by the cDNA may be used to screen several

CC molecules or compounds to identify a ligand that specifically binds to

CC the protein, or to produce or purify an antibody to the protein that can

CC be used to detect a protein in a sample or purify a natural or

CC recombinant protein from a sample. The nucleotides may be useful for

CC diagnosing, staging, treating, or monitoring the progression of treatment

CC of a vascular disease, e.g. atherosclerosis, cancer, coronary artery

CC disease, hypertension, diabetes, pre-eclampsia, ischaemia-reperfusion

CC injury, restenosis, or stroke. The cDNAs can also be used for large-scale

CC genetic or gene expression analysis of several new nucleic acid

CC molecules. Antibodies to the proteins encoded by the cDNAs are useful for

CC diagnosing pre-pathologic disorders, and chronic or acute diseases

CC associated with abnormalities in the expression, amount or distribution

CC of the protein. The present sequence represents a cDNA of the invention

CC that is differentially expressed in activated vascular tissue. Note: The

CC sequence data for this patent did not form part of the specification, but

CC was obtained in electronic format directly from USPTO at

XX <http://seqdata.uspto.gov/sequence.html?DocID=20020137081>

XX

XX Sequence 7132 BP; 1702 A; 1879 C; 1942 G; 1609 T; 0 U; 0 Other;

XX

Query Match 10.7%; Score 18; DB 8; Length 7132;

Best Local Similarity 100.0%; Pred. No. 12;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX

QY 121 TTCCAATGCTGGTGTGAA 138

Db 259 TTCCAATGCTGGTGTGAA 276

XX

RESULT 13

ID ADD93418

XX ADD93418 standard; cDNA; 7973 BP.

XX

AC ADD93418;

XX

XX 29-JAN-2004 (first entry)

XX

XX Human lipid-associated molecule LIPAM-6 polynucleotide.

XX

XX Human; lipid-associated molecule; LIPAM-6; neuroprotective; relaxant;

XX antithyroid; antidiabetic; cytostatic; dermatological; immunosuppressive;

XX antiinflammatory; thymimetic; antiallergic; cerebroprotective;

XX gastrointestinal; hepatotropic; nephrotropic; anticonvulsant;

XX antiparkinsonian; antibacterial; antiparasitic; fungicide; protozoacide;

XX virucide; uropathic; antirheumatic; cardiant; cardiovascular; anti-HIV;

XX nootropic; gene; ss.

XX

```

OS XX Homo sapiens
XX FH Key Location/Qualifiers
XX FT CDS 138..5858
XX FT /*tag= b
XX FT sig_peptide 138..197
XX FT /*tag= a
XX FT /*product= "Human LIPAM-6"
XX FT mat_peptide 198..5855
XX FT /*tag= c
XX PN WO2003083081-A2.
XX XX 09-OCT-2003.
XX XX 27-MAR-2003; 2003WO-US009755.
XX XX 29-MAR-2002; 2002US-0368722P.
XX XX 03-MAY-2002; 2002US-0377576P.
XX XX 05-JUL-2002; 2002US-0393934P.
XX XX 27-SEP-2002; 2002US-041269P.
XX XX (INCY-) INCYTE CORP.
XX XX Emerling BM, Marquis JP, Chawla NK, Lee SY, Duggan BM, Warren BA;
XX XX Baughn MR, Lee EA, Griffin JA, Kahle AE, Elliott VS, Chang H;
XX XX Lee S, Ramkumar J, Bulloch SA, Hafalia AJA, Khare R, Jiang X;
XX XX Jackson AA;
XX XX WPI; 2003-788347/74.
XX XX P-PSDB; ADD93399.
XX XX New LIPAM polypeptides, useful for diagnosing, preventing, and treating
XX XX disorders associated with abnormal expression or activity of LIPAM, e.g.
XX XX neuromuscular, immunological, cardiovascular disorders, cancer and/or
XX XX infections.
XX XX Claim 80; Page 222-224; 238pp; English.
XX XX The present sequence is the nucleotide sequence of human lipid-associated
XX XX molecule LIPAM-6 (incyte polynucleotide 7510885CBI), which encodes a
XX XX protein that shows sequence homology to chicken alpha-2-macroglobulin
XX XX receptor. This is one of 19 LIPAM polynucleotides of the invention. The
XX XX invention relates to novel LIPAMs and the nucleic acids encoding them,
XX XX and to the use of nucleic acids and proteins in the diagnosis, treatment
XX XX and prevention of disorders associated with abnormal expression or
XX XX activity of LIPAM such as neurodegenerative disorders (e.g. Parkinson's
XX XX disease, Alzheimer's disease), muscular disorders (e.g. myotonic
XX XX dystrophy, catantonia), endocrine disorders (e.g. diabetes, Grave's
XX XX disease), cancers (e.g. leukaemia, cervical or breast cancers),
XX XX immunological disorders (e.g. scleroderma, systemic lupus erythematosus,
XX XX allergies), gastrointestinal disorders (e.g. Crohn's disease), renal
XX XX disorders (e.g. Goodpasture's syndrome), infections (e.g. viral,
XX XX bacterial, fungal, parasitic, protozoal, helminthic), cardiovascular
XX XX disorders (e.g. atherosclerosis), or hepatic diseases (e.g. cirrhosis).
XX XX The invention also relates to the assessment of the effects of exogenous
XX XX compounds on the expression of nucleic acids and LIPAMs. The invention
XX XX provides expression vectors, host cells, antibodies, agonists and
XX XX antagonists, transgenic organisms, and arrays and microarrays of the
XX XX polynucleotides.
XX XX Sequence 7973 BP; 1849 A; 2140 C; 2244 G; 1740 T; 0 U; 0 Other;
XX XX Query Match 10.7%; Score 18; DB 10; Length 7973;
XX XX Best Local Similarity 100.0%; Pred. No. 12;
XX XX Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX XX Qy 121 TTCCAATGCTGGTGTGAA 138
XX XX Db 1386 TTCCAATGCTGGTGTGAA 1403
XX XX RESULT 14

```

```

AAD47370
ID AAD47370 standard; DNA; 7974 BP.
XX AC AAD47370;
XX DT 24-FEB-2003 (first entry)
XX DE Human LP288 DNA.
XX XX Human; LP protein; cell proliferative disorder; actinic keratinosis;
XX KW arteriosclerosis; psoriasis; leukaemia; lymphoma; autoimmune disorder;
XX KW melanoma; cancer; inflammatory disorder; Addison's disease; allergy;
XX KW acquired immune deficiency syndrome; AIDS; ankylosing spondylitis;
XX KW amyloidosis; anaemia; asthma; Crohn's disease; Goodpasture's syndrome;
XX KW gout; Grave's disease; Hashimoto's thyroiditis; cardiovascular disorder;
XX KW congestive heart failure; ischaemic heart disease; myocardial infarction;
XX KW angina pectoris; atherosclerosis; hypertension; neurological disorder;
XX KW stroke; Parkinson's disease; Alzheimer's disease; developmental disorder;
XX KW Down's syndrome; cerebral palsy; gene therapy; nephrotropic; nontropic;
XX KW gene; ds.
XX OS Homo sapiens.
XX XX Key Location/Qualifiers
XX XX CDS 142..5859
XX XX /*tag= a
XX XX /*product= "Human LP288 protein"
XX XX sig_peptide 142..201
XX XX /*tag= b
XX XX mat_peptide 202..2856
XX XX /*tag= c
XX XX /*product= "Human mature LP288 protein"
XX XX WO200274906-A2.
XX XX 26-SEP-2002.
XX XX 01-MAR-2002; 2002WO-US005093.
XX XX 16-MAR-2001; 2001US-0276596P.
XX XX 13-APR-2001; 2001US-0283654P.
XX XX 20-APR-2001; 2001US-0285238P.
XX XX 03-MAY-2001; 2001US-0288548P.
XX XX 11-MAY-2001; 2001US-0290351P.
XX XX (ELIL ) LILLY & CO ELI.
XX XX Amegadzie BY, Basinski MB, Chen D, Huang C, Keleher GP,
XX XX Perkins DR, Rosteck PRJ, Rowlinson SW, Sankharam PR, Seno ET;
XX XX Su EW, Zhi Y;
XX XX WPI; 2003-018798/01.
XX XX P-PSDB; AAE29923.
XX XX New mammalian LP polynucleotides and proteins, useful in gene therapy,
XX XX e.g. for treating or preventing cancers, autoimmune (e.g. AIDS),
XX XX cardiovascular (e.g. myocardial infarction) or neurological (e.g. stroke)
XX XX disorders.
XX XX Claim 3; Page 235-243; 290pp; English.
XX XX The present invention relates to LP (LP318a, LP318b, LP288, LP289, LP343,
XX XX LP319a, LP319b, LP321, LP317, LP283, LP344, LP345 or LP346) proteins and
XX XX polynucleotides encoding such proteins. Sequences of the invention are
XX XX used to diagnose, treat or prevent cell proliferative disorders (e.g.
XX XX actinic keratinosis, arteriosclerosis, psoriasis, leukaemia, lymphoma,
XX XX melanoma, brain cancer or breast cancer), autoimmune or inflammatory
XX XX disorders (e.g. AIDS (acquired immune deficiency syndrome), Addison's
XX XX disease, allergies, ankylosing spondylitis, amyloidosis, anaemia, asthma,
XX XX Crohn's disease, Goodpasture's syndrome, gout, Hashimoto's thyroiditis or
XX XX Grave's disease), cardiovascular disorders (e.g. congestive heart
XX XX failure, ischaemic heart disease, angina pectoris, myocardial infarction,
XX XX atherosclerosis or hypertension), neurological disorders (e.g. stroke,

```

CC Parkinson's disease or Alzheimer's disease) or developmental disorders
 CC (e.g. Down's syndrome or cerebral palsy). They are also used in gene
 CC therapy. The present sequence is human LP288 DNA
 XX
 SQ Sequence 7974 BP; 1848 A; 2143 C; 2243 G; 1740 T; 0 U; 0 Other;
 Query Match 10.7%; Score 18; DB 8; Length 7974;
 Best Local Similarity 100.0%; Pred. No. 12;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 121 TTCCAATGCTGCTGAA 138
 |||||
 DB 1387 TTCCAATGCTGCTGAA 1404
 |||||
 RESULT 15
 AAF94902/c
 ID AAF94902 standard; cDNA; 396 BP.
 XX
 AC AAF94902;
 XX
 DT 23-MAY-2001 (first entry)
 XX
 DE Human ovarian cancer associated coding sequence SEQ ID NO: 93.
 XX
 KW Human, ovarian cancer; vaccine; gene therapy; carcinoma; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200118046-A2.
 XX
 PD 15-MAR-2001.
 XX
 PF 08-SEP-2000; 2000WO-US024827.
 XX
 PR 10-SEP-1999; 99US-00394374.
 PR 01-MAY-2000; 2000US-00561778.
 PR 15-AUG-2000; 2000US-00640173.
 PR 07-SEP-2000; 2000US-00656668.
 XX
 PA (CORI-) CORIXA CORP.
 XX
 PI Xu J, Stolk JA;
 XX
 DR WPI; 2001-211395/21.
 XX
 PT Isolated polypeptides associated with ovarian carcinomas, and the nucleic
 PT acids that encode them, useful for the prevention diagnosis and treatment
 PT of ovarian cancers.
 XX
 PS Claim 5; Page 146; 189pp; English.
 XX
 CC The present invention provides a number of coding sequences and proteins,
 CC the over-expression of which is associated with ovarian carcinoma/cancer.
 CC These can be used in the diagnosis, treatment and prevention of ovarian
 CC cancer, optionally by gene therapy or in the form of a vaccine. The
 CC present sequence is an example of one of these sequences
 XX
 SQ Sequence 396 BP; 110 A; 86 C; 79 G; 112 T; 0 U; 9 Other;
 Query Match 10.1%; Score 17; DB 4; Length 396;
 Best Local Similarity 100.0%; Pred. No. 46;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 140 TTCTGAGGATGAGAAC 156
 |||||
 DB 199 TTCTGAGGATGAGAAC 183
 |||||

Search completed: August 26, 2005, 21:18:51
 Job time : 317.243 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 26, 2005, 18:47:02 ; Search time 1855.79 Seconds
(without alignments)
4386.526 Million cell updates/sec

Title: US-10-617-978-14_COPY_73_240
Perfect score: 168
Sequence: 1 ccgggaactaccacttga.....atgsaacgtgaaggtctga 168

Scoring table: OLIGO NUC
Gapop_60.0 , Gapext 60.0

Searched: 4708233 seqs, 24227607955 residues

Word size : 0

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : GenEmbl.*

- 1: gb.ba.*
- 2: gb.htg.*
- 3: gb.in.*
- 4: gb.on.*
- 5: gb.ov.*
- 6: gb.pat.*
- 7: gb.ph.*
- 8: gb.pl.*
- 9: gb.pr.*
- 10: gb.ro.*
- 11: gb.sts.*
- 12: gb.sy.*
- 13: gb.un.*
- 14: gb.vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	22	13.1	126889	10	AC141867	AC141867 Mus muscu
c 2	22	13.1	213178	10	AC148021	AC148021 Mus muscu
3	22	13.1	234325	2	AC133845	AC133845 Rattus no
4	22	13.1	233305	2	AC094603	AC094603 Rattus no
c 5	22	13.1	246625	2	AC094293	AC094293 Rattus no
c 6	21	12.5	205947	2	AC117346	AC117346 Rattus no
c 7	21	12.5	291547	2	AC114722	AC114722 Rattus no
c 8	20	11.9	45550	2	AC015339	AC015339 Drosophil
c 9	20	11.9	169680	3	AC008337	AC008337 Drosophil
c 10	20	11.9	183048	3	AC012376	AC012376 Drosophil
c 11	20	11.9	207871	2	AC119359	AC119359 Rattus no
c 12	20	11.9	297970	3	AE003511	AE003511 Drosophil
c 13	19	11.3	184568	9	AC009558	AC009558 Homo sapi
c 14	19	11.3	186716	2	AC025843	AC025843 Homo sapi
c 15	19	11.3	190414	10	AC150648	AC150648 Mus muscu
c 16	19	11.3	199810	10	AC125050	AC125050 Mus muscu
c 17	19	11.3	301448	2	AC097707	AC097707 Rattus no
c 18	18	10.7	192	3	AY351304	AY351304 Centruroi
c 19	18	10.7	192	3	AY351305	AY351305 Centruroi

20	18	10.7	192	3	AY351306	AY351306 Centruroi
21	18	10.7	192	3	AY351307	AY351307 Centruroi
22	18	10.7	319	3	AF338459	AF338459 Centruroi
23	18	10.7	320	3	AF338461	AF338461 Centruroi
24	18	10.7	323	3	AF338454	AF338454 Centruroi
25	18	10.7	323	3	AF338455	AF338455 Centruroi
26	18	10.7	323	3	AF338456	AF338456 Centruroi
27	18	10.7	323	3	AF338457	AF338457 Centruroi
28	18	10.7	323	3	AF338458	AF338458 Centruroi
29	18	10.7	323	3	AF338462	AF338462 Centruroi
30	18	10.7	323	3	AF338463	AF338463 Centruroi
31	18	10.7	323	3	AF338464	AF338464 Centruroi
32	18	10.7	819	5	AY553186	AY553186 Xenopus 1
33	18	10.7	899	5	AB077363	AB077363 Xenopus 1
34	18	10.7	940	8	AY051046	AY051046 Arabidops
35	18	10.7	1206	5	BC056846	BC056846 Xenopus 1
36	18	10.7	1209	5	BC045084	BC045084 Xenopus 1
37	18	10.7	1867	8	AK175540	AK175540 Arabidops
38	18	10.7	1989	8	AF360290	AF360290 Arabidops
c 39	18	10.7	3045	8	AB089697	AB089697 Fagopyrum
c 40	18	10.7	3061	8	AB089696	AB089696 Fagopyrum
41	18	10.7	5185	6	CQ717587	CQ717587 Sequence
42	18	10.7	5587	6	AX921066	AX921066 Sequence
43	18	10.7	6846	9	AB011540	AB011540 Homo sapi
c 44	18	10.7	41601	9	AC004017	AC004017 Homo sapi
c 45	18	10.7	100627	9	AC007359	AC007359 Homo sapi

ALIGNMENTS

RESULT 1	AC141867	126889 bp	DNA	linear	ROD 27-NOV-2003
LOCUS	AC141867				
DEFINITION	Mus musculus BAC clone RP24-424E17	from chromosome 19,	complete		
ACCESSION	AC141867				
VERSION	AC141867.4	GI:33667259			
KEYWORDS	HTG.				
SOURCE	Mus musculus (house mouse)				
ORGANISM	Mus musculus				
REFERENCE	1 (bases 1 to 126889)				
AUTHORS	Shah, N.				
TITLE	The sequence of Mus musculus BAC clone RP24-424E17				
JOURNAL	Unpublished (2001)				
REFERENCE	2 (bases 1 to 126889)				
AUTHORS	Wilson, R.				
TITLE	Sequencing of Mus musculus				
JOURNAL	Unpublished (2001)				
REFERENCE	3 (bases 1 to 126889)				
AUTHORS	McPherson, J.D. and Waterston, R.H.				
TITLE	Direct Submission				
JOURNAL	Submitted (19-MAR-2003) Genome Sequencing Center, 4444 Forest Park				
REFERENCE	4 (bases 1 to 126889)				
AUTHORS	Wilson, R.K.				
TITLE	Direct Submission				
JOURNAL	Submitted (03-JUL-2003) Genome Sequencing Center, 4444 Forest Park				
REFERENCE	5 (bases 1 to 126889)				
AUTHORS	Wilson, R.K.				
TITLE	Direct Submission				
JOURNAL	Submitted (15-AUG-2003) Genome Sequencing Center, 4444 Forest Park				
REFERENCE	6 (bases 1 to 126889)				
AUTHORS	Wilson, R.				
TITLE	Direct Submission				
JOURNAL	Submitted (27-NOV-2003) Department of Genetics, Washington				
REFERENCE	On Aug 15, 2003 this sequence version replaced gi:32441365.				
COMMENT	----- Genome Center				

Center: Washington University Genome Sequencing Center

Center code: WUGSC

Web site: <http://genome.wustl.edu>

Contact: submissions@watson.wustl.edu

----- Summary Statistics

----- Center project name: M_BB0424E17

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. Wes Warren, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu>

SOURCE INFORMATION:

The RPCI-24 BAC Library has been constructed by Pieter de Jong and coworkers (<http://www.chori.org>) from male C57BL/6J mouse spleen and/or brain genomic DNA. The clone and detailed information can be obtained from Pieter de Jong and coworkers at <http://www.chori.org>

NEIGHBORING SEQUENCE INFORMATION:

This sequence is the entire insert of the clone.

FEATURES
source

1..126889
Location/Qualifiers
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="19"
/map="19"
/clone="RP24-424E17"
/clone_lib="RPCI-24"

repeat_region	61..144	/rpt family="B4"	/rpt family="Alu"
repeat_region	475..587	/rpt family="B4"	/rpt family="ERV1"
repeat_region	1530..1769	/rpt family="B4"	/rpt family="B4"
repeat_region	1851..2309	/rpt family="MER2_type"	/rpt family="B4"
repeat_region	2328..2592	/rpt family="MER2_type"	/rpt family="B4"
repeat_region	2667..2750	/rpt family="B4"	/rpt family="B4"
repeat_region	2787..2839	/rpt family="MER2_type"	/rpt family="Alu"
repeat_region	3019..3196	/rpt family="ERV1"	/rpt family="B4"
repeat_region	4329..4553	/rpt family="L1"	/rpt family="B4"
repeat_region	4630..4748	/rpt family="B2"	/rpt family="B4"
repeat_region	4644..4804	/rpt family="B4"	/rpt family="B4"
repeat_region	5359..5438	/rpt family="MaLR"	/rpt family="B4"
repeat_region	5528..5696	/rpt family="MER1_type"	/rpt family="B4"
repeat_region	6674..6957	/rpt family="B2"	/rpt family="B4"
repeat_region	7441..7592	/rpt family="B4"	/rpt family="B4"

```

repeat_region 43836..44005
                /rpt_family="B4"
repeat_region 44316..44536
                /rpt_family="B2"
repeat_region 44782..44985
                /rpt_family="MER1_type"
repeat_region 46558..46675
                /rpt_family="L1"
repeat_region 46738..47143
                /rpt_family="L1"
repeat_region 47216..47495
                /rpt_family="L1"
repeat_region 47530..47595
                /rpt_family="ERV1"

```

```

Query Match      13.1%; Score 22; DB 10; Length 126889;
Best Local Similarity 100.0%; Pred. No. 0.2;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 72 CATTAAAGATCTGTGAGAAACAC 93
Db 47185 CATTAAAGATCTGTGAGAAACAC 47206

```

```

RESULT 2
AC148021/c
LOCUS          AC148021          213178 bp      DNA      linear      ROD 29-MAY-2004
DEFINITION    Mus musculus BAC clone RP23-9A7 from chromosome 19, complete
               sequence.
ACCESSION     AC148021
VERSION       AC148021.2  GI:47825191
KEYWORDS      HTG.
SOURCE        Mus musculus
               Mus musculus (house mouse)
ORGANISM      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE     1 (bases 1 to 213178)
AUTHORS       Levy, A., Cotton, M., Haglund, K. and Bielicki, L.
TITLE         The sequence of Mus musculus BAC clone RP23-9A7
JOURNAL       Unpublished (2001)
REFERENCE     2 (bases 1 to 213178)
AUTHORS       Wilson, R.K.
TITLE         Direct Submission
JOURNAL       Submitted (28-JAN-2004) Genome Sequencing Center, 4444 Forest Park
               Parkway, St. Louis, MO 63108, USA
REFERENCE     3 (bases 1 to 213178)
AUTHORS       Wilson, R.K.
TITLE         Direct Submission
JOURNAL       Submitted (29-MAY-2004) Genome Sequencing Center, Washington
               University School of Medicine, 4444 Forest Park Parkway, St. Louis,
               MO 63108, USA
COMMENT       On May 29, 2004 this sequence version replaced gi:41351753.
               ----- Genome Center
               Center: Washington University Genome Sequencing Center
               Center code: WUGSC
               Web site: http://genome.wustl.edu
               Contact: submissions@watson.wustl.edu
               ----- Summary Statistics
               -----
               Center project name: M_BA0009A07

```

NOTICE:

This sequence was finished as follows unless otherwise noted:
 all regions were double stranded, sequenced with an alternate
 chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such
 as compressions and repeats; all regions were covered by sequence
 from more than one subclone; and the assembly was confirmed by
 restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. Wes Warren.

Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu>

SOURCE INFORMATION:

The RPCI-23 BAC Library has been constructed by Kazutoyo Oseawa and Minako Tateni in the laboratory of Pieter de Jong (<http://www.chori.org>) from female C57BL/6J mouse kidney and/or brain genomic DNA. The clone and detailed information can be obtained from Research Genetics, Inc. (<http://www.reagen.com>) or Pieter de Jong and coworkers at <http://www.chori.org>

NEIGHBORING SEQUENCE INFORMATION:

This sequence is the entire insert of the clone. This clone is overlapped by AC134830 and AC140378.

FEATURES	Location/Qualifiers
source	1..213178 /organism="Mus musculus" /mol_type="genomic DNA" /db_xref="taxon:10090" /chromosome="19" /map="19" /clone="RP23-9A7" /clone_lib="RPCI-23" 797..899
repeat_region	/rpt_family="L1" 1041..1292
repeat_region	/rpt_family="B4" 1520..1924
repeat_region	/rpt_family="RMER17C" 1708..1972
repeat_region	/rpt_family="ERVK" 2113..2728
repeat_region	/rpt_family="L1" 3189..3235
repeat_region	/rpt_family="ERV1" 4224..4582
repeat_region	/rpt_family="ERV1" 4583..4804
repeat_region	/rpt_family="MER1_type" 4805..4850
repeat_region	/rpt_family="ERV1" 4851..5023
repeat_region	/rpt_family="MaLR" 5024..5056
repeat_region	/rpt_family="ERV1" 5880..6128
repeat_region	/rpt_family="L1" 6270..6641
repeat_region	/rpt_family="L1" 6834..7274
repeat_region	/rpt_family="L1" 7408..8060
repeat_region	/rpt_family="L1" 8132..8261
repeat_region	/rpt_family="Alu" 8260..8562
repeat_region	/rpt_family="MaLR" 8262..8304
repeat_region	/rpt_family="B4" 9073..9310
repeat_region	/rpt_family="L1" 9831..9903
repeat_region	/rpt_family="ERVK" 14939..15004
repeat_region	/rpt_family="ID" 19319..19610
repeat_region	/rpt_family="MaLR" 23410..23607
repeat_region	/rpt_family="MER1_type" 24037..24073
repeat_region	/rpt_family="B2" 24221..24317

```
repeat_region 58332 .58389 /rpt_family="L2"
repeat_region 61467 .61619 /rpt_family="L2"
repeat_region 61959 .62140 /rpt_family="L1"
repeat_region 62379 .62642 /rpt_family="L1"

Query Match 13.1%; Score 22; DB 10; Length 213178;
Best Local Similarity 100.0%; Pred. No. 0.2;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 72 CATTAAGATCTGTCAGAAACAC 93
Db 101861 CATTAAGATCTGTCAGAAACAC 101840

RESULT 3
AC133845
LOCUS AC133845
DEFINITION Rattus norvegicus clone CH230-41J2, *** SEQUENCING IN PROGRESS ***,
2 unordered pieces.
ACCESSION AC133845
VERSION AC133845.3 GI:30522110
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE
1 (bases 1 to 234325)
Muzny, D., Marie, Metzker, M., Lee, Abramson, S., Adams, C., Alder, J.,
Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D.,
Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,
Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,
Cardenas, V., Carter, K., Cavazos, I., Caesar, H., Center, A.,
Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,
Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
Davila, M., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,
Egan, A., Escoto, M., Eugene, C., Evans, C., Falls, T., Fan, G.,
Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
Fraser, C., Gabis, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
Georgiev, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W.,
Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K.,
Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,
Hernandez, R., Hines, S., Hladun, S., Hodgson, A., Hogue, M.,
Hollins, B., Howells, S., Hulyk, S., Hume, J., Jolebird, D., Jackson, A.,
Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,
Karpach, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,
Kowis, C., Kraft, C., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,
Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,
Lorensu, L., Lousegh, H., Lozano, R., Lu, X., Ma, J.,
Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A.,
Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,
Mawhiney, S., McLeod, M., McNeill, T., Meenen, E.,
Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,
Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L.,
Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,
Nwaokeme, O., Okwuonu, G., Olarnpunsagoon, A., Pal, S., Parks, K.,
Pasternak, S., Paul, H., Perez, A., Perez, L., Prankoch, C.,
Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.,
Puzo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M., Reigh, R.,
Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,
Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.,
Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H.,
Shetty, J., Shvartsbeyan, A., Sisson, I., Sitter, C., Smales, D.,
Sneed, A., Sodergren, E., Song, X., Sorelle, R., Sosa, J.,
Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C.,
Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K.,
```

Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J., Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F., Williams,G., Willson,R., Wleczkyk,R., Wooden,H., Worley,K., Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V., Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O., Weinstock,G. and Gibbs,R.A.

Direct Submission

TITLE

JOURNAL

REFERENCE

AUTHORS

JOURNAL

Submitted (19-SEP-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

REFERENCE

AUTHORS

JOURNAL

Submitted (10-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

COMMENT

On May 10, 2003 this sequence version replaced gi:25013050. The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: GSVF

Center clone name: CH230-41J2

----- Summary Statistics

Assembly program: Atlas 3.0;

Consensus quality: 224086 bases at least Q40

Consensus quality: 226523 bases at least Q30

Consensus quality: 227917 bases at least Q20

Estimated insert size: 237878; sum-of-contigs estimation

Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 2 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

* 1 231823: contig of 231823 bp in length

* 231824 231923: gap of unknown length

* 231924 234325: contig of 2402 bp in length.

----- Location/Qualifiers

1. .234325

/organism="Rattus norvegicus"

/mol_type="genomic DNA"

/db_xref="taxon:10116"

/clone="CH230-41J2"

1. .3168

/note="wgs_end extension

clone end:Sp6"

6731. .53657

FEATURES

source

misc_feature

misc_feature

/note="clone boundary
 clone_end:Sp6
 site:ECORI
 end_sequence:BH342505"
 complement(230261..231118)
 /note="clone boundary
 clone_end:T7
 site:ECORI
 end_sequence:BH342503"

misc_feature

ORIGIN

Query Match 13.1%; Score 22; DB 2; Length 234325;
 Best Local Similarity 100.0%; Pred. No. 0.2;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 67 GACTGCATTAGATCTGTCTGACA 88
 |||||

Db 54847 GACTGCATTAGATCTGTCTGACA 54868
 |||||

RESULT 4

AC094603

LOCUS

DEFINITION

AC094603

VERSION

KEYWORDS

SOURCE

ORGANISM

Rattus norvegicus

Rattus norvegicus

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

Rattus.

REFERENCE

AUTHORS

Murny,D,Warie., Metzker,M,Lea., Abramzon,S., Adams,C., Alder,J.,

Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,

Anyalebechi,V., Ayoyagi,A., Ayodeji,M., Baca,E., Baden,H.,

Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,

Biewalt,N., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,

Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,

Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,

Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,X., Chen,Z., Chu,J.,

Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,

Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,

Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,

Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,

Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,

Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,

Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,

Gebregorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,

Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,

Harvey,Y., Havlak,P., Haves,A., Henderson,N., Hernandez,J.,

Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,

Hollins,B., Howells,S., Huylk,S., Hume,J., Idlebird,D., Jackson,A.,

Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,

Karpachy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,

Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,

Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,

Lorensuhwa,L., Loulseghe,H., Lozado,R.J., Lu,X., Ma,J.,

Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,

Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,

Mawhney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,

Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,

Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,

Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,

Nwaokemele,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K.,

Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C.,

Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L.,

Puzo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,

Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,

Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,

Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,

Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajda,D.,

Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,

Tattum,B., Tattum,B., Tattum,B., Tattum,B., Tattum,B.,

Tattum,B., Tattum,B., Tattum,B., Tattum,B., Tattum,B.,

Tattum,B., Tattum,B., Tattum,B., Tattum,B., Tattum,B.,

Tattum,B., Tattum,B., Tattum,B., Tattum,B., Tattum,B.,

Tattum,B., Tattum,B., Tattum,B., Tattum,B., Tattum,B.,

Tattum,B., Tattum,B., Tattum,B., Tattum,B., Tattum,B.,

Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K., Valas,R., Vera,V., Villasana,B., Waldron,L., Walker,B., Wang,J., Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F., Williams,G., Willson,R., Wleczek,R., Woodson,H., Worley,K., Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V., Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O., Weinstock,G. and Gibbs,R.A.

TITLE
JOURNAL
REFERENCE
AUTHORS
JOURNAL

2 (bases 1 to 246625)
Worley,K.C.
Direct Submission
Submitted (17-SEP-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

REFERENCE
AUTHORS
TITLE
JOURNAL

3 (bases 1 to 246625)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (09-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

COMMENT

On May 9, 2003 this sequence version replaced gi:25188904.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'), within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GAIN
Center clone name: CH230-3K24
----- Summary Statistics
Assembly program: Atlas
Consensus quality: 184728 bases at least Q40
Consensus quality: 191694 bases at least Q30
Consensus quality: 196691 bases at least Q20
Estimated insert size: 187878; sum-of-contigs estimation
Quality coverage: 5x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 20 contigs. The true order of the pieces
* is, not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

* 1 205043: contig of 205043 bp in length
* 205044
* 205144
* 215060: contig of 9917 bp in length
* 215061
* 215160: gap of unknown length
* 215161
* 216255: contig of 1095 bp in length
* 216356
* 216355: gap of unknown length
* 216356
* 217394: contig of 1039 bp in length
* 217395
* 217494: gap of unknown length
* 217495
* 218555: contig of 1061 bp in length
* 218556
* 218555: gap of unknown length
* 218656
* 220239: contig of 1584 bp in length
* 220240
* 220339: gap of unknown length

* 220340
* 222257: contig of 1918 bp in length
* 222258
* 222357: gap of unknown length
* 223852: contig of 1495 bp in length
* 223853
* 223952: gap of unknown length
* 223953
* 224660: contig of 1508 bp in length
* 225461
* 225560: gap of unknown length
* 225561
* 226912: contig of 1352 bp in length
* 226913
* 227012: gap of unknown length
* 227013
* 228059: contig of 1047 bp in length
* 228060
* 228159: gap of unknown length
* 228160
* 229457: contig of 1297 bp in length
* 229457
* 229556: gap of unknown length
* 229557
* 230918: contig of 1362 bp in length
* 230919
* 231018: gap of unknown length
* 231019
* 232608: contig of 1590 bp in length
* 232609
* 232708: gap of unknown length
* 232709
* 233944: contig of 1236 bp in length
* 233945
* 234044: gap of unknown length
* 234045
* 237241: contig of 3197 bp in length
* 237242
* 237341: gap of unknown length
* 238817
* 238818
* 238917: gap of unknown length
* 238918
* 240337: contig of 1420 bp in length
* 240338
* 240437: gap of unknown length
* 240438
* 242451: contig of 2014 bp in length
* 242452
* 242551: gap of unknown length
* 242552
* 246625: contig of 4074 bp in length.

FEATURES
source
1. .246625
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-3K24"
113870..115506
/note="wgs_contig"
157302..158702
/note="wgs_contig"
188107..190360
/note="wgs_contig"
213846..215060
/note="wgs_contig"

ORIGIN

Query Match 13.1%; Score 22; DB 2; Length 246625;
Best Local Similarity 100.0%; Pred. No. 0.2;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 72 CATTAGATCTGTCTCAGAAACAC 93
|||||
Db 65928 CATTAGATCTGTCTCAGAAACAC 65907

RESULT 6
AC117346/c
LOCUS
DEFINITION Rattus norvegicus clone CH230-391F21, WORKING DRAFT SEQUENCE, 3
unordered pieces.
AC117346
AC117346
AC117346
AC117346.6 GI:24635507
HTG; HTGS PHASE1; HTGS DRAFT; HTGS_FULLTOP.
KEYWORDS
Rattus norvegicus (Norway rat)
SOURCE
ORGANISM
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE
1 (bases 1 to 205947)
AUTHORS
Muzny,D.,Marie., Metzker,M.,Lee., Abramson,S., Adams,C., Alder,J.,
Allen,C., Allen,H., Alsbrooks,S., Amin,A., Angiano,D.,
Ayalebechi,V.,Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
Bliswal,D., Blair,J., Blankenburg,K., Blych,P., Brown,M.,
Bryant,N., Buhay,C., Burrell,K., Burrell,K., Calderon,E.,
Cardenas,V., Carter,K., Cavazos,I., Caesar,H., Center,A.,

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 1 (bases 1 to 291547)
 Muzny D, Marie, Metzker, M. Lee., Abramson, S., Adams, C., Alder, J.,
 Allen, C., Allen, H., Albrooks, S., Amin, A., Anguiano, D.,
 Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,
 Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
 Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
 Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderton, E.,
 Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A.,
 Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,
 Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
 Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
 Delgado, O., Denison, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
 Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,
 Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G.,
 Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
 Fraser, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
 Gebregregoris, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W.,
 Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K.,
 Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,
 Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hogue, M.,
 Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, P., Jackson, A.,
 Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,
 Karpach, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,
 Kowis, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,
 Liu, U., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,
 Lorehuahua, L., Loulsegred, H., Lozano, R. J., Lu, X., Ma, J.,
 Maheshwari, M., Mahindaratne, M., Mahmoud, M., Malloy, K., Mangum, A.,
 Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, B.,
 Mawhinney, S., McLeod, M. P., McNeill, T. Z., Meenen, E.,
 Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,
 Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L.,
 Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,
 Nwokedemeh, O., Okwuonu, G., Olarnpunsagoon, A., Pal, S., Parks, K.,
 Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C.,
 Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.-L.,
 Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R.,
 Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,
 Rives, C., Rodkey, T., Rojao, A., Rose, M., Rose, R., Ruiz, S. J.,
 Sanders, W., Savary, G., Scherer, S., Scott, G., Shatsman, S., Shen, H.,
 Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C. D., Smaiz, D.,
 Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J. D.,
 Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C.,
 Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K.,
 Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J.,
 Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,
 Williams, G., Willson, R., Wleczyk, R., Wooden, H., Worley, K.,
 Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,
 Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von
 Niederhausern, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O.,
 Weinstein, G. and Gibbs, R. A.

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu

----- Project Information
 Center project name: GTJR
 Center clone name: CH230-60K20

----- Summary Statistics
 Assembly program: Phrap; version 0.990329
 Consensus quality: 231275 bases at least Q40
 Consensus quality: 234311 bases at least Q30
 Consensus quality: 235939 bases at least Q20

 Estimated insert size: 240124; sum-of-contigs estimation
 Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 1 contigs. Gaps between the contigs
 * are represented as runs of N. The order of the pieces
 * is believed to be correct as given, however the sizes
 * of the gaps between them are based on estimates that have
 * provided by the submitter.
 * This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.

* 1 291547: contig of 291547 bp in length.

FEATURES

source

1 .291547
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10116"
 /clone="CH230-60K20"

misc_feature

1 .1044
 /notes="wgs_end_extension
 clone_end:Sp6"
 complement(8119..8720)
 /notes="clone_boundary
 clone_end:Sp6
 site:"

misc_feature

end sequence:BH365931"
 21921..119834
 /notes="clone_boundary
 clone_end:T7
 site:"

misc_feature

end sequence:BH365929"
 32443..34180
 /notes="wgs_end_extension
 clone_end:T7"

ORIGIN

Query Match 12.5%; Score 21; DB 2; Length 291547;
 Best Local Similarity 100.0%; Pred. No. 0.74;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 135 TGAATTTCTGAGGATGAGAA 155

Db 13561 TGAATTTCTGAGGATGAGAA 13541

RESULT 8

AC015339/c

LOCUS

DEFINITION

AC015339

AC015339

VERSION

AC015339.1 GI:6435996
 45550 bp DNA linear HTG 16-NOV-1999
 Drosophila melanogaster, *** SEQUENCING IN PROGRESS ***.

Submitted (19-NOV-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA

On Nov 19, 2002 this sequence version replaced gi:23265675.

The sequence in this assembly is a combination of BAC based reads
 and whole genome shotgun sequencing reads assembled using Atlas
 (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described
 in the feature table below represents a scaffold in the Atlas
 assembly (a 'contig-scaffold'). Within each contig-scaffold,
 individual sequence contigs are ordered and oriented, and separated


```

/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/strain="Y; cn bw sp"
/db_xref="taxon:7227"
/chromosome="X"
/map="188-18B"
/clone="BAC clone BACR48C12 (D1122)"
/clone_lib="RPCI-98 (Roswell Park Cancer Institute
Drosophila melanogaster BAC library, partial EcoRI in
pBACs3.6)"

```

ORIGIN

Query Match 11.9%; Score 20; DB 3; Length 183048;
 Best Local Similarity 100.0%; Pred. No. 2.9;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 118 GCCTTCCCAATGCTGCTGTA 137
 Db 66140 GCCTTCCCAATGCTGCTGTA 66121

RESULT 11

AC119359 207871 bp DNA linear HTG 20-NOV-2002
 LOCUS Rattus norvegicus clone CH230-466N16, WORKING DRAFT SEQUENCE, 2
 UNORDERED PAGES.

ACCESSION AC119359
 VERSION AC119359.5 GI:25137741
 KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
 SOURCE Rattus norvegicus (Norway rat)

ORGANISM

Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.

REFERENCE

1 (bases 1 to 207871)
 Murny,D.Marie., Metzker,M.Lee., Abramson,S., Adams,C., Alder,J.,
 Allen,C., Allen,H., Albrooks,S., Amin,A., Anguiano,D.,
 Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
 Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
 Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
 Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
 Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
 Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
 Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
 Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
 Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
 Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
 Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
 Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
 Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
 Gebréorgis,E., Gear,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
 Gunaratne,P., Haaland,W., Hamill,C., Hamilton,C., Hamilton,K.,
 Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
 Hernández,R., Hines,S., Hladun,S.L., Hodgson,A., Hogue,M.,
 Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
 Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
 Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovac,C.,
 Kowib,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
 Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
 Lu,X., Ma,J.,
 Maheshwara,L., Louissegh,H., Lozado,R.J., Lu,X., Ma,J.,
 Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
 Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
 Mayhew,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
 Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
 Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
 Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
 Nwackelmech,O., Okwou,G., Olarnpunsagoon,A., Pal,S., Parks,K.,
 Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkuch,C.,
 Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L.,
 Puzo,M., Quiroz,J., Rachin,E., Reeves,K., Regier,M.A., Reigh,R.,
 Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
 Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
 Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,

FEATURES
 source

1. 207871
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10116"
 /clone="CH230-466N16"

Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D.,
 Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
 Steimle,M., Strong,R., Sutton,A., Svatek,A., Taber,P., Taylor,C.,
 Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
 Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,
 Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
 Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K.,
 Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
 Yu,P., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von
 Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
 Weinstein,G. and Gibbs,R.A.
 Direct Submission
 Unpublished
 2 (bases 1 to 207871)
 Worley,K.C.
 Direct Submission
 Submitted (26-APR-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 207871)
 Rat Genome Sequencing Consortium.
 Direct Submission
 Submitted (20-NOV-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 On Nov 20, 2002 this sequence version replaced gi:23908428.
 The sequence in this assembly is a combination of BAC based reads
 and whole genome shotgun sequencing reads assembled using Atlas
 (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
 in the feature table below represents a scaffold in the Atlas
 assembly (a 'contig-scaffold'). Within each contig-scaffold,
 individual sequence contigs are ordered and oriented, and separated
 by sized gaps filled with Ns to the estimated size. The sequence
 may extend beyond the ends of the clone and there may be sequence
 contigs within a contig-scaffold that consist entirely of whole
 genome shotgun sequence reads. Both end sequences and whole genome
 shotgun sequence only contigs will be indicated in the feature
 table.

----- Genome Center
 Center: Baylor College of Medicine

Center code: BCM

Web site: http://www.hgsc.bcm.tmc.edu/

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: GVFX

Center clone name: CH230-466N16

----- Summary Statistics

Assembly program: Phrap; version 0.990329

Consensus quality: 177278 bases at least Q40

Consensus quality: 180313 bases at least Q30

Consensus quality: 181910 bases at least Q20

Estimated insert size: 183823; sum-of-contigs estimation

Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
 (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 consists of 2 contigs. The true order of the pieces
 is not known and their order in this sequence record is
 arbitrary. Gaps between the contigs are represented as
 runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 as soon as it is available and the accession number will
 be preserved.

* 1 206612: contig of 206612 bp in length

* 206613 206712: gap of unknown length

* 206713 207871: contig of 1159 bp in length.

Location/Qualifiers

1. 207871
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10116"
 /clone="CH230-466N16"

```

misc_feature      1. 1047
                  /note="wgs_end_extension
clone_end:Sp6"
misc_feature      24305. 25227
                  /note="clone_boundary
clone_end:Sp6
site:
end_sequence:B2216496"
misc_feature      45252. 46059
                  /note="clone_boundary
clone_end:T7
site:
end_sequence:B2216494"
misc_feature      205108. 206612
                  /note="wgs_end_extension
clone_end:T7"

ORIGIN
Query Match      11.9%; Score 20; DB 2; Length 207871;
Best Local Similarity 100.0%; Pred. No. 2.9;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 132 GTGTGAATTTCTGAAGGATG 151
      |||||
Db 118039 GTGTGAATTTCTGAAGGATG 118058

RESULT 12
AE003511/c
LOCUS
DEFINITION Drosophila melanogaster chromosome X, section 63 of 74 of the
complete sequence.
ACCESSION AE003511
VERSION AE003511.3 Gi:22832537
KEYWORDS
SOURCE
ORGANISM Drosophila melanogaster (fruit fly)
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Ephydroidea; Drosophila; Drosophila.
REFERENCE 1 (bases 1 to 297970)
          Adams,M.D., Celniker,S.E., Holt,R.A., Evans,C.A., Gocayne,J.D.,
          Ananidis,P.G., Scherer,S.E., Li,P.W., Hoekins,R.A., Galle,R.F.,
          Sutton,G.G., Wortman,J.R., Yandell,M.D., Zhang,Q., Chen,L.X.,
          Brandon,R.C., Rogers,Y.H., Blazej,R.G., Champe,M., Pfeiffer,B.D.,
          Wan,K.H., Doyle,C., Baxter,E.G., Helt,G., Nelson,C.R., Gabor,G.L.,
          Abail,J.F., Agbayani,A., An,H.J., Andrews-Pfannkoch,C., Baldwin,D.,
          Ballou,R.M., Basu,A., Baxendale,J., Bayraktaroglu,L., Beasley,E.M.,
          Beeson,K.Y., Benos,P.V., Berman,B.P., Bhandari,D., Bolshakov,S.,
          Borkova,D., Botchan,M.R., Bouck,J., Brokstein,P., Brotter,P.,
          Burtis,K.C., Busam,D.A., Butler,H., Cadieu,E., Center,A.,
          Chandra,I., Cherry,J.M., Cawley,S., Dahlke,C., Davenport,L.B.,
          Davies,P., de Pablos,B., Delcher,A., Deng,Z., Mays,A.D., Dew,I.,
          Dietz,S.M., Dodson,K., Doup,L.E., Downes,M., Dugan-Rocha,S.,
          Dunkov,B.C., Dunn,P., Durbin,K.J., Evangelist,C.C., Ferrar,C.,
          Ferrieri,S., Fleischmann,W., Fosler,K., Gabriellian,A.E., Garg,N.S.,
          Gelbart,W.M., Glasser,K., Glodek,A., Gong,F., Gorrell,J.H., Gu,Z.,
          Guan,P., Harris,M., Harris,N.L., Harvey,D., Heiman,T.J.,
          Hernandez,J.R., Houck,J., Hostin,D., Houston,K.A., Howland,T.J.,
          Wei,M.H., Ibegwam,C., Jalali,M., Kalush,F., Karpen,G.H., Ke,Z.,
          Kennison,J.A., Ketchum,K.A., Kimmel,B.E., Kodira,C.D., Kraft,C.,
          Kravitz,S., Kulp,D., Lai,Z., Lasko,P., Lei,Y., Levitsky,A.A.,
          Li,J., Li,Z., Liang,Y., Lin,X., Liu,X., Mattel,B., McIntosh,T.C.,
          McLeod,M.P., McPherson,D., Merkulov,G., Milshina,N.V., Moberly,C.,
          Morris,J., Moshrefi,A., Mount,S.M., Moy,M., Murphy,B., Murphy,L.,
          Muzny,D.M., Nelson,D.L., Nelson,D.R., Nelson,K.A., Nixon,K.,
          Nuskern,D.R., Pacleb,J.M., Palazzolo,M., Pitman,G.S., Pan,S.,
          Pollard,J., Puri,V., Reese,M.C., Reinert,K., Remington,K.,
          Saunders,R.D., Scheeler,F., Shen,H., Shue,B.C., Siden-Kiamos,I.,
          Simpson,M., Skupski,M.P., Smith,T., Spier,E., Spradling,A.C.,
          Stapleton,M., Strong,R., Sun,E., Svirska,R., Tector,C., Turner,R.,
          Venter,E., Wang,A.H., Wang,X., Wang,Z.Y., Wassarman,D.A.,

```

```

Weinstock,G.M., Weissenbach,J., Williams,S.M., Woodage,T.,
Worley,K.C., Wu,D., Yang,S., Yao,Q.A., Ye,J., Yeh,R.F.,
Zaveri,J.S., Zhan,M., Zhao,G., Zhao,Q., Zheng,L., Zheng,X.H.,
Zhong,F.N., Zhong,W., Zhou,X., Zhu,S., Zhu,X., Smith,H.O.,
Gibbs,R.A., Myers,E.W., Rubin,G.M. and Venter,J.C.
The genome sequence of Drosophila melanogaster
Science 287 (5461), 2185-2195 (2000)
20196006
10731132
2 (bases 1 to 297970)
Celniker,S.E., Wheeler,D.A., Kronmiller,B., Carlson,J.W.,
Halpern,A., Patel,S., Adams,M., Champe,M., Dugan,S.P., Frisoe,E.,
Hodgson,A., George,R.A., Hoskins,R.A., Laverty,T., Muzny,D.M.,
Nelson,C.R., Pacleb,J.M., Park,S., Pfeiffer,B.D., Richards,S.,
Sodergren,E.J., Svirska,R., Tabor,P.E., Wan,K., Stapleton,M.,
Sutton,G.G., Venter,C., Weinstock,G., Scherer,S.E., Myers,E.W.,
Gibbs,R.A. and Rubin,G.M.
Finishing a whole-genome shotgun: release 3 of the Drosophila
melanogaster euchromatic genome sequence
Genome Biol. 3 (12), RESEARCH0079 (2002)
22426065
12537568
3 (bases 1 to 297970)
Misra,S., Crosby,M.A., Mungall,C.J., Matthews,B.B., Campbell,K.S.,
Hradecky,P., Huang,Y., Kaminker,J.S., Millburn,G.H., Prochnik,S.E.,
Smith,C.D., Tupy,J.L., Whitfield,E.J., Bayraktaroglu,L.,
Berman,B.P., Bettencourt,B.R., Celniker,S.E., de Grey,A.D.,
Drysdaile,R.A., Harris,N.L., Richter,J., Russo,S., Schroeder,A.J.,
Shu,S.Q., Stapleton,M., Yamada,C., Ashburner,M., Gelbart,W.M.,
Rubin,G.M. and Lewis,S.E.
Annotation of the Drosophila melanogaster euchromatic genome: a
systematic review
Genome Biol. 3 (12), RESEARCH0083 (2002)
22426069
12537572
4 (bases 1 to 297970)
Kaminker,J.S., Bergman,C.M., Kronmiller,B., Carlson,J.,
Svirska,R., Patel,S., Frisoe,E., Wheeler,D.A., Lewis,S.E.,
Rubin,G.M., Ashburner,M. and Celniker,S.E.
The transposable elements of the Drosophila melanogaster
euchromatin: a genomics perspective
Genome Biol. 3 (12), RESEARCH0084 (2002)
22426070
12537573
5 (bases 1 to 297970)
Adams,M.D., Celniker,S.E., Gibbs,R.A., Rubin,G.M. and Venter,C.J.
Direct Submission
Submitted (21-MAR-2000) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
6 (bases 1 to 297970)
FlyBase
Direct Submission
Submitted (06-SEP-2002) University of California Berkeley, 539 Life
Sciences Addition, Berkeley, CA 94720, USA
7 (bases 1 to 297970)
FlyBase
Direct Submission
Submitted (10-MAR-2004) FlyBase, Harvard University, Biological
Laboratories, 16 Divinity Avenue, Cambridge, MA 02138, USA
On Sep 13, 2002 this sequence version replaced gi:10728334.
Location/Qualifiers
1. 297970
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/chromosome="X"
complement(5294..10011)
/locus_tag="TE19084"
/map="18A3-18A3"
/transposon="Doc{}}163"
/db_xref="FLYBASE:FBti10019084"
repeat_region

```



```

Theodore,J., Travers,M., Travis,N., Trigilio,J., Vassiliev H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zemdek,D., Zimmer,A. and Zody,W.
TITLE      Direct Submission
JOURNAL    Submitted (31-MAR-2001) Whitehead Institute/MIT Center for Genome
COMMENT    Research, 320 Charles Street, Cambridge, MA 02141, USA
On Mar 4, 2001 this sequence version replaced gi:13176254.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L1678
Center clone name: 142_J_21
-----
FEATURES             Location/Qualifiers
     source            1. .184568
                        /organism="Homo sapiens"
                        /mol_type="genomic DNA"
                        /db_xref="taxon:9606"
                        /chromosome="15"
                        /map="15"
                        /clones="RP11-142J21"
                        /clone_lib="REC1-11 Human Male BAC"
repeat_region        1095..1130
                        /rpt_family="TAA)n"
repeat_region        complement(1133..1353)
                        /rpt_family="MIR"
repeat_region        complement(1469..1590)
                        /rpt_family="MER45A"
repeat_region        complement(1791..2156)
                        /rpt_family="THE1C"
repeat_region        3640..3936
                        /rpt_family="AluSx"
repeat_region        3937..4023
                        /rpt_family="GA-rich"
repeat_region        4299..4424
                        /rpt_family="MER91C"
repeat_region        complement(4597..4895)
                        /rpt_family="L2"
repeat_region        5032..5245
                        /rpt_family="MIR"
repeat_region        5586..5606
                        /rpt_family="AT-rich"
repeat_region        5715..5769
                        /rpt_family="CA)n"
repeat_region        5771..5796
                        /rpt_family="GA)n"
repeat_region        complement(6505..6675)
                        /rpt_family="AluYa5"
repeat_region        7057..7222
                        /rpt_family="L1ME3"
repeat_region        complement(7943..8117)
                        /rpt_family="MIR"
repeat_region        8287..8573
                        /rpt_family="AluSx"
repeat_region        8606..8922
                        /rpt_family="AluSx"
repeat_region        complement(9851..10096)
                        /rpt_family="MIR"
repeat_region        16256..16556
                        /rpt_family="AluSx"
repeat_region        complement(17636..17701)
                        /rpt_family="L2"
repeat_region        17665..17799
                        /rpt_family="AT-rich"
repeat_region        complement(17807..18096)
                        /rpt_family="AluSp"
repeat_region        20270..20562
                        /rpt_family="AluSx"
repeat_region        complement(20829..21099)
                        /rpt_family="AluSx"
repeat_region        complement(21609..21770)
                        /rpt_family="MIR"
repeat_region        22857..22887
                        /rpt_family="(T)n"
repeat_region        23088..23107
                        /rpt_family="(T)n"
repeat_region        23719..23767
                        /rpt_family="AT-rich"
repeat_region        23913..24170
                        /rpt_family="MIR"
repeat_region        24174..24180
                        /note="Less than 30 qual SNGL region"
repeat_region        24190..24245
                        /note="single clone coverage"
repeat_region        24225..24245
                        /note="Less than 30 qual SNGL region"
repeat_region        24267..24300
                        /note="Less than 30 qual SNGL region"
repeat_region        complement(24549..24830)
                        /rpt_family="LTR16B"
repeat_region        complement(25294..25371)
                        /rpt_family="MIR"
repeat_region        25407..25629
                        /rpt_family="MIR"
repeat_region        25768..26030
                        /rpt_family="MIR"
repeat_region        26335..26616
                        /rpt_family="Tigger7"
repeat_region        26830..26711
                        /rpt_family="Tigger7"
repeat_region        27162..27341
                        /rpt_family="MIR"
repeat_region        27496..27586
                        /rpt_family="LIME"
repeat_region        27755..28032
                        /rpt_family="L2"
repeat_region        complement(28102..28244)
                        /rpt_family="MLT1F"
repeat_region        complement(28446..28559)
                        /rpt_family="MLT1I"
repeat_region        28565..28691
                        /rpt_family="L2"
repeat_region        complement(29285..29352)
                        /rpt_family="MIR"
repeat_region        complement(30157..30298)
                        /rpt_family="MER105"
repeat_region        30298..30808
                        /rpt_family="MLT1D"
repeat_region        complement(30809..30865)
                        /rpt_family="MER105"
repeat_region        complement(31945..32109)
                        /rpt_family="MIR"
repeat_region        33917..34293
                        /rpt_family="MLT1A2"
repeat_region        complement(34302..34856)
                        /rpt_family="MLT2A2"
repeat_region        complement(34857..34952)
                        /rpt_family="MER45"
repeat_region        complement(34959..35090)
                        /rpt_family="MER45"
repeat_region        complement(35133..35401)
                        /rpt_family="L2"
repeat_region        35403..35532
                        /rpt_family="L2"
repeat_region        36899..37000
                        /rpt_family="MIR"
repeat_region        complement(37203..37325)
                        /rpt_family="MIR"
repeat_region        38250..38545
                        /rpt_family="AluSx"

```

```

repeat_region complement(38562. .38626)
/rpt family="MIR"
repeat_region 38955. .39209

```

```

Query Match 11.3%; Score 19; DB 9; Length 184568;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Oy 133 TGTGAATTCCTGAAGGATG 151
      |||||
Db 51960 TGTGAATTCCTGAAGGATG 51978

```

```

RESULT 14
AC025843
LOCUS AC025843 186716 bp DNA linear HTG 06-MAY-2001
DEFINITION Homo sapiens chromosome 15 clone RP11-313A9 map 15, WORKING DRAFT
SEQUENCE, 18 unordered pieces.
ACCESSION AC025843
VERSION AC025843.3 GI:8076815
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 186716)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Ahabraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,H.,
Bozulavskiy,L., Boukhgalter,B., Brown,A., Burkett,G.,
Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
Collins,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S.,
Dodge,S., Domino,M., Doyle,M., Ferreira,P., Fitzhugh,W., Gage,D.,
Galagan,J., Gardyna,S., Glende,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
Kleih,J., Lacroque,K., Lamazares,R., Landers,T., Lehoczyk,J.,
Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,
McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,
Melchior,J., Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J.,
Murphy,T., Navlor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
O'Neil,D., Oliver,T.M., Oliver,J., Peterson,K., Pierre,N.,
Piachi,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,
Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
Young,G., Zainoun,J., Zimmer,A. and Zody,M.
Direct Submission
Submitted (16-MAR-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On May 25, 2000 this sequence version replaced gi:7408040.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

```

```

----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: 16179
Center clone name: 313 A_9
----- Summary Statistics
Sequencing vector: M13, W7815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 175296 bases at least Q40
Consensus quality: 180588 bases at least Q30
Consensus quality: 183068 bases at least Q20

```

```

TITLE
JOURNAL
COMMENT
Submitted (16-MAR-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On May 25, 2000 this sequence version replaced gi:7408040.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

```

```

----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: 16179
Center clone name: 313 A_9
----- Summary Statistics
Sequencing vector: M13, W7815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 175296 bases at least Q40
Consensus quality: 180588 bases at least Q30
Consensus quality: 183068 bases at least Q20

```

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

```

Insert size: 188000; agarose-fp
Insert size: 185016; sum-of-contigs
Quality coverage: 4.4 in Q20 bases; agarose-fp
Quality coverage: 4.4 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 18 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

```

```

* 1 1251: contig of 1251 bp in length
* 1251: gap of 100 bp
* 1352 4014: contig of 2663 bp in length
* 4015 5680: contig of 1566 bp in length
* 5681 7344: contig of 100 bp
* 7345 8320: contig of 1564 bp in length
* 8321 8420: gap of 100 bp
* 8421 9927: contig of 876 bp in length
* 9928 10027: gap of 100 bp
* 10028 12018: contig of 1507 bp in length
* 12019 12118: contig of 1991 bp in length
* 12119 15745: contig of 100 bp
* 15746 15845: contig of 3627 bp in length
* 15846 17954: gap of 100 bp
* 17955 28155: contig of 3908 bp in length
* 28156 28255: gap of 100 bp
* 28256 36313: contig of 8064 bp in length
* 36314 47349: gap of 100 bp
* 47350 47445: contig of 10930 bp in length
* 47446 61276: contig of 100 bp
* 61277 61376: contig of 13827 bp in length
* 61377 80216: contig of 100 bp
* 80217 80316: gap of 100 bp
* 80317 101592: contig of 21276 bp in length
* 101593 101692: gap of 100 bp
* 101693 125692: contig of 24000 bp in length
* 125693 125792: gap of 100 bp
* 125793 151472: contig of 25680 bp in length
* 151473 151572: gap of 100 bp
* 151573 186716: contig of 35144 bp in length.

```

FEATURES

source

```

1. 186716
   /organism="Homo sapiens"
   /mol_type="Genomic DNA"
   /db_xref="taxon:9606"
   /chromosome="15"
   /map="15"
   /clone="RP11-313A9"
   /clone_lib="RPC1-11 Human Male BAC"
1. 1251
   /note="assembly_fragment"
1352. 4014
   /note="assembly_fragment"
4115. 5680
   /note="assembly_fragment"
5781. 7344
   /note="assembly_fragment"
7445. 8320
   /note="assembly_fragment"
clone end:SP6
vector_side:left
8421. 9927
   /note="assembly_fragment"
10028. 12018
   /note="assembly_fragment"

```



```
misc_feature 12119..15745
/note="assembly_fragment"
misc_feature 15846..19753
/note="assembly_fragment"
misc_feature 19854..28155
/note="assembly_fragment"
misc_feature 28256..36319
/note="assembly_fragment
clone_end:17
vector_side:left"
misc_feature 36420..47349
/note="assembly_fragment"
misc_feature 47450..61276
/note="assembly_fragment"
misc_feature 61377..80216
/note="assembly_fragment"
misc_feature 80317..101592
/note="assembly_fragment"
misc_feature 101693..125692
/note="assembly_fragment"
misc_feature 125793..151472
/note="assembly_fragment"
misc_feature 151573..186716
/note="assembly_fragment"

ORIGIN
Query Match 11.3%; Score 19; DB 2; Length 186716;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 133 TGTGAATTTCTGAAGGATG 151
|||||
Db 131167 TGTGAATTTCTGAAGGATG 131185

RESULT 15
AC150648/c
LOCUS AC150648 190414 bp DNA linear ROD 15-SEP-2004
DEFINITION Mus musculus chromosome 7 clone RP23-235N5, complete sequence.
ACCESSION AC150648
VERSION AC150648.2 GI:52077975
KEYWORDS HTG.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Wilson,R.K.
TITLE The sequence of Mus musculus clone
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 190414)
AUTHORS Wilson,R.K.
TITLE Direct Submission
JOURNAL Submitted (29-JUL-2004) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
REFERENCE 3 (bases 1 to 190414)
AUTHORS Wilson,R.K.
TITLE Direct Submission
JOURNAL Submitted (15-SEP-2004) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
COMMENT On Sep 15, 2004 this sequence version replaced gi:50811813.

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu
Contact: submissions@wustl.edu
----- Project Information -----
Center project name: M_BA0235N05
-----
Location/Qualifiers
1..190414
/organism="Mus musculus"

/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="7"
/clone="RP23-235N5"

ORIGIN
Query Match 11.3%; Score 19; DB 10; Length 190414;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 118 GCCTTCCAATGCTGGTGTG 136
|||||
Db 106866 GCCTTCCAATGCTGGTGTG 106848

Search completed: August 26, 2005, 22:22:31
Job time : 1864.79 secs
```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 26, 2005, 18:47:02 ; Search time 1955.21 Seconds
(without alignments)
4386.526 Million cell updates/sec

Title: US-10-617-978-14_COPY_64_240

Perfect score: 177
Sequence: 1 gctgactcccggaacta.....atgagaactgaaggtctga 177

Scoring table:
OLIGO NUC
Gapop 60.0 , Gapext 60.0

Searched: 4708233 seqs, 24227607955 residues

Word size : 0

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

GenEmbl.*

1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: gb.vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	22	12.4	126889	10 AC141867	AC141867 Mus muscu
2	22	12.4	213178	10 AC148021	AC148021 Mus muscu
3	22	12.4	234325	2 AC133845	AC133845 Rattus no
4	22	12.4	235305	2 AC094603	AC094603 Rattus no
5	22	12.4	246625	2 AC094293	AC094293 Rattus no
6	21	11.9	205947	2 AC117346	AC117346 Rattus no
7	21	11.9	291547	2 AC114722	AC114722 Rattus no
8	20	11.3	45550	2 AC015339	AC015339 Drosophil
9	20	11.3	169680	3 AC008337	AC008337 Drosophil
10	20	11.3	183048	3 AC012376	AC012376 Drosophil
11	20	11.3	207871	2 AC119359	AC119359 Rattus no
12	20	11.3	297970	3 AB003511	AB003511 Drosophil
13	19	10.7	184568	9 AC009558	AC009558 Homo sapi
14	19	10.7	186716	2 AC025843	AC025843 Homo sapi
15	19	10.7	190414	10 AC150648	AC150648 Mus muscu
16	19	10.7	199810	10 AC125050	AC125050 Mus muscu
17	19	10.7	301448	2 AC097707	AC097707 Rattus no
18	18	10.2	192	3 AY351304	AY351304 Centruroi
19	18	10.2	192	3 AY351305	AY351305 Centruroi

20	18	10.2	192	3	AY351306	AY351306 Centruroi
21	18	10.2	192	3	AY351307	AY351307 Centruroi
22	18	10.2	319	3	AF338459	AF338459 Centruroi
23	18	10.2	320	3	AF338461	AF338461 Centruroi
24	18	10.2	323	3	AF338454	AF338454 Centruroi
25	18	10.2	323	3	AF338455	AF338455 Centruroi
26	18	10.2	323	3	AF338456	AF338456 Centruroi
27	18	10.2	323	3	AF338457	AF338457 Centruroi
28	18	10.2	323	3	AF338458	AF338458 Centruroi
29	18	10.2	323	3	AF338462	AF338462 Centruroi
30	18	10.2	323	3	AF338463	AF338463 Centruroi
31	18	10.2	323	3	AF338464	AF338464 Centruroi
32	18	10.2	819	5	AY553186	AY553186 Xenopus 1
33	18	10.2	899	5	AB077363	AB077363 Xenopus 1
34	18	10.2	940	8	AY051046	AY051046 Arabidops
35	18	10.2	1206	5	BC056846	BC056846 Xenopus 1
36	18	10.2	1209	5	BC045084	BC045084 Xenopus 1
37	18	10.2	1867	8	AK175540	AK175540 Arabidops
38	18	10.2	1989	8	AF360290	AF360290 Arabidops
C 39	18	10.2	3045	8	AB089697	AB089697 Fagopyrum
C 40	18	10.2	3061	8	AB089696	AB089696 Fagopyrum
41	18	10.2	5185	6	CQ717587	CQ717587 Sequence
42	18	10.2	5587	6	AX921066	AX921066 Sequence
43	18	10.2	6846	9	AB011540	AB011540 Homo sapi
C 44	18	10.2	41601	9	AC004017	AC004017 Homo sapi
C 45	18	10.2	100627	9	AC007359	AC007359 Homo sapi

ALIGNMENTS

RESULT 1	AC141867	126889 bp	DNA	linear	ROD 27-NOV-2003
LOCUS	Mus musculus BAC clone RP24-424E17 from chromosome 19, complete sequence.				
DEFINITION	AC141867.4 GI:33667259				
ACCESSION	AC141867				
VERSION	HTG.				
KEYWORDS	Mus musculus (house mouse)				
SOURCE	Mus musculus				
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
REFERENCE	Shah, N.				
AUTHORS	The sequence of Mus musculus BAC clone RP24-424E17				
JOURNAL	Unpublished (2001)				
REFERENCE	2 (bases 1 to 126889)				
AUTHORS	Wilson, R.				
TITLE	Sequencing of Mus musculus				
JOURNAL	Unpublished (2001)				
REFERENCE	3 (bases 1 to 126889)				
AUTHORS	McPherson, J.D. and Waterston, R.H.				
TITLE	Direct Submission				
JOURNAL	Submitted (19-MAR-2003) Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO 63108, USA				
REFERENCE	4 (bases 1 to 126889)				
AUTHORS	Wilson, R.K.				
TITLE	Direct Submission				
JOURNAL	Submitted (03-JUL-2003) Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO 63108, USA				
REFERENCE	5 (bases 1 to 126889)				
AUTHORS	Wilson, R.K.				
TITLE	Direct Submission				
JOURNAL	Submitted (15-AUG-2003) Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO 63108, USA				
REFERENCE	6 (bases 1 to 126889)				
AUTHORS	Wilson, R.				
TITLE	Direct Submission				
JOURNAL	Submitted (27-NOV-2003) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA				
COMMENT	On Aug 15, 2003 this sequence version replaced gi:32441365.				

Center: Washington University Genome Sequencing Center
 Center code: WUGSC
 Web site: <http://genome.wustl.edu>
 Contact: submissions@watson.wustl.edu
 ----- Summary Statistics
 Center project name: M_BB0424E17

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. Wes Warren, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu>

SOURCE INFORMATION:

The RPCI-24 BAC Library has been constructed by Pieter de Jong and coworkers (<http://www.chori.org>) from male C57BL/6J mouse spleen and/or brain genomic DNA. The clone and detailed information can be obtained from Pieter de Jong and coworkers at <http://www.chori.org>

NEIGHBORING SEQUENCE INFORMATION:

This sequence is the entire insert of the clone.

FEATURES

source	
1. .126889	/organsims="Mus musculus"
	/mol_type="genomic DNA"
	/db_xref="taxon:10090"
	/chromosome="19"
	/map="19"
	/clone="RP24-424E17"
	/clone.lib="RPCI-24"
61. .144	/rpt_family="B4"
475. .587	/rpt_family="Achoho"
1530. .1769	/rpt_family="MER2_type"
1851. .2309	/rpt_family="MER2_type"
2328. .2592	/rpt_family="B4"
2667. .2750	/rpt_family="MER2_type"
2787. .2839	/rpt_family="ERV1"
3019. .3196	/rpt_family="L1"
4329. .4553	/rpt_family="B2"
4630. .4748	/rpt_family="B4"
4644. .4804	/rpt_family="MaLR"
5359. .5438	/rpt_family="MER1_type"
5528. .5636	/rpt_family="B2"
6674. .6957	/rpt_family="B4"
7441. .7592	
repeat_region	/rpt_family="Alu"
7733. .7797	/rpt_family="ERV1"
7852. .7988	/rpt_family="B4"
10219. .10319	/rpt_family="L1"
10430. .10636	/rpt_family="B4"
11101. .11238	/rpt_family="Alu"
11436. .11648	/rpt_family="B2"
12403. .12535	/rpt_family="MIR"
13942. .14026	/rpt_family="L2"
14863. .14986	/rpt_family="MaLR"
15804. .15851	/rpt_family="L2"
15907. .16111	/rpt_family="L2"
16318. .16364	/rpt_family="L2"
16376. .16475	/rpt_family="MaLR"
16658. .16841	/rpt_family="B2"
18345. .18444	/rpt_family="L2"
19872. .19990	/rpt_family="MIR"
21585. .21911	/rpt_family="MaLR"
23518. .23593	/rpt_family="ID"
24449. .24567	/rpt_family="MIR"
26603. .26709	/rpt_family="Alu"
27720. .27775	/rpt_family="Mariner"
28356. .28401	/rpt_family="Mariner"
31094. .31241	/rpt_family="Alu"
31323. .31599	/rpt_family="B4"
32264. .32339	/rpt_family="ID"
32394. .32559	/rpt_family="B2"
32585. .32731	/rpt_family="Alu"
32867. .33064	/rpt_family="MER1_type"
35162. .35332	/notes="Unresolved simple sequence repeat."
35419. .35658	/rpt_family="L1"
37357. .37569	/rpt_family="B2"
38521. .38763	/rpt_family="MER2_type"
39563. .39768	/rpt_family="B2"
40172. .40385	/rpt_family="L1"
40410. .40683	/rpt_family="L1"
43617. .43835	/rpt_family="B4"

```

repeat_region 43836..44005
/rpt_family="B4"
repeat_region 44316..44536
/rpt_family="B2"
repeat_region 44782..44985
/rpt_family="MER1_type"
repeat_region 46558..46675
/rpt_family="L1"
repeat_region 46738..47143
/rpt_family="L1"
repeat_region 47216..47495
/rpt_family="L1"
repeat_region 47530..47595
/rpt_family="ERV1"

```

Query Match 12.4%; Score 22; DB 10; Length 126889;

Best Local Similarity 100.0%; Pred.No. 0.24; Mismatches 0; Indels 0; Gaps 0;

```

Qy 81 CATTAAAGATCTGTCAGAAACAC 102
|||||
Db 47185 CATTAAAGATCTGTCAGAAACAC 47206
|||||

```

```

RESULT 2
AC148021/c 213178 bp DNA linear ROD 29-MAY-2004
LOCUS Mus musculus BAC clone RP23-9A7 from chromosome 19, complete
DEFINITION sequence.
ACCESSION AC148021
VERSION AC148021.2 GI:47825191
KEYWORDS HTG.
SOURCE Mus musculus (house mouse)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 213178)
AUTHORS Levy,A., Cotton,M., Haglund,K. and Bielecki,L.
TITLE The Sequence of Mus musculus BAC clone RP23-9A7
JOURNAL Unpublished (2001)
REFERENCE 2 (bases 1 to 213178)
AUTHORS Wilson,R.K.
TITLE Direct Submission
JOURNAL Submitted (28-JAN-2004) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
REFERENCE 3 (bases 1 to 213178)
AUTHORS Wilson,R.K.
TITLE Direct Submission
JOURNAL Submitted (29-MAY-2004) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
COMMENT On May 29, 2004 this sequence version replaced gi:41351753:
----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu
Contact: submissions@watson.wustl.edu
----- Summary Statistics
-----

```

NOTICE:

This sequence was finished as follows unless otherwise noted:
 all regions were double stranded, sequenced with an alternate
 chemistry, or covered by high quality data (i.e., phred quality >=
 30); an attempt was made to resolve all sequencing problems, such
 as compressions and repeats; all regions were covered by sequence
 from more than one subclone; and the assembly was confirmed by
 restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. Wes Warren,

Department of Genetics, Washington University, St. Louis MO. For
 additional information about the map position of this sequence, see
 http://genome.wustl.edu

SOURCE INFORMATION:

The RPCI-23 BAC Library has been constructed by Kazutoyo Oseawa
 and Minako Tatenio in the laboratory of Pieter de Jong
 (http://www.chori.org) from female C57BL/6J mouse kidney and/or
 brain genomic DNA. The clone and detailed information can be
 obtained from Research Genetics, Inc. (http://www.resgen.com) or
 Pieter de Jong and coworkers at http://www.chori.org

NEIGHBORING SEQUENCE INFORMATION:

This sequence is the entire insert of the clone. This clone is
 overlapped by AC134830 and AC140378.

FEATURES	Location/Qualifiers
Source	1..213178
	/organism="Mus musculus"
	/mol_type="genomic DNA"
	/db_xref="taxon:10090"
	/chromosome="19"
	/map="19"
	/clone="RP23-9A7"
	/clone_lib="RPCI-23"
	797..899
repeat_region	/rpt_family="L1"
	1041..1292
repeat_region	/rpt_family="B4"
	1520..1924
repeat_region	/rpt_family="RMER17C"
	1708..1972
repeat_region	/rpt_family="ERVK"
	2113..2728
repeat_region	/rpt_family="L1"
	3189..3235
repeat_region	/rpt_family="ERV1"
	4224..4582
repeat_region	/rpt_family="ERV1"
	4583..4804
repeat_region	/rpt_family="MER1_type"
	4805..4850
repeat_region	/rpt_family="ERV1"
	4851..5023
repeat_region	/rpt_family="MaLR"
	5024..5056
repeat_region	/rpt_family="ERV1"
	5880..6128
repeat_region	/rpt_family="L1"
	6270..6641
repeat_region	/rpt_family="L1"
	6834..7274
repeat_region	/rpt_family="L1"
	7408..8060
repeat_region	/rpt_family="L1"
	8132..8261
repeat_region	/rpt_family="Alu"
	8260..8562
repeat_region	/rpt_family="MaLR"
	8262..8304
repeat_region	/rpt_family="B4"
	9073..9310
repeat_region	/rpt_family="L1"
	9831..9903
repeat_region	/rpt_family="ERVK"
	14939..15004
repeat_region	/rpt_family="ID"
	19319..19610
repeat_region	/rpt_family="MaLR"
	23410..23607
repeat_region	/rpt_family="MER1_type"
	24037..24073
repeat_region	/rpt_family="B2"
	24221..24317

```
repeat_region /rpt family="L2"
24495. .24580 /rpt family="L2"
repeat_region /rpt family="L2"
27452. .27603 /rpt family="L1"
repeat_region /rpt family="L1"
27734. .27880 /rpt family="L1"
repeat_region /rpt family="L1"
28012. .28200 /rpt family="MIR"
repeat_region /rpt family="MIR"
29405. .29566 /rpt family="MIR"
repeat_region /rpt family="MIR"
30402. .30644 /rpt family="L1"
repeat_region /rpt family="L1"
30772. .30865 /rpt family="L1"
repeat_region /rpt family="L1"
30868. .31157 /rpt family="MaLR"
repeat_region /rpt family="MaLR"
31226. .31343 /rpt family="L1"
repeat_region /rpt family="L1"
31571. .31856 /rpt family="L1"
repeat_region /rpt family="L1"
31883. .32147 /rpt family="B4"
repeat_region /rpt family="B4"
32323. .32744 /rpt family="L1"
repeat_region /rpt family="L1"
32776. .33125 /rpt family="MaLR"
repeat_region /rpt family="MaLR"
33140. .33245 /rpt family="Alu"
repeat_region /rpt family="Alu"
33418. .33823 /rpt family="L1"
repeat_region /rpt family="L1"
34901. .34987 /rpt family="Alu"
repeat_region /rpt family="Alu"
35612. .35711 /rpt family="MER1_type"
repeat_region /rpt family="MER1_type"
36132. .36218 /rpt family="L1"
repeat_region /rpt family="L1"
36252. .36298 /rpt family="MaLR"
repeat_region /rpt family="MaLR"
36299. .36429 /rpt family="Alu"
repeat_region /rpt family="Alu"
39089. .39164 /rpt family="Alu"
repeat_region /rpt family="Alu"
39565. .39643 /rpt family="MaLR"
repeat_region /rpt family="MaLR"
40187. .40547 /rpt family="MaLR"
repeat_region /rpt family="MaLR"
42281. .42318 /rpt family="Alu"
repeat_region /rpt family="Alu"
43201. .43361 /rpt family="Alu"
repeat_region /rpt family="Alu"
45578. .45698 /rpt family="Alu"
repeat_region /rpt family="Alu"
46623. .46752 /rpt family="MER1_type"
repeat_region /rpt family="MER1_type"
50496. .50683 /rpt family="L1"
repeat_region /rpt family="L1"
50808. .50992 /rpt family="L1"
repeat_region /rpt family="L1"
51082. .51239 /rpt family="L1"
repeat_region /rpt family="L1"
51240. .51434 /rpt family="B2"
repeat_region /rpt family="B2"
51435. .51537 /rpt family="L1"
repeat_region /rpt family="L1"
51550. .51683 /rpt family="Alu"
repeat_region /rpt family="Alu"
52489. .52851 /rpt family="MaLR"
repeat_region /rpt family="MaLR"
54229. .54373 /rpt family="Alu"
repeat_region /rpt family="Alu"
57514. .57861 /rpt family="MaLR"
```

```
repeat_region 58332. .58389 /rpt family="B2"
repeat_region 61467. .61619 /rpt family="L1"
repeat_region 61959. .62140 /rpt family="L1"
repeat_region 62379. .62642 /rpt family="L1"

Query Match 12.4%; Score 22; DB 10; Length 213178;
Best Local Similarity 100.0%; Pred. No. 0.23;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 81 CATTAGATCTGTCAGAAACAC 102
Db 101861 CATTAGATCTGTCAGAAACAC 101840

RESULT 3
AC133845
LOCUS AC133845
DEFINITION Rattus norvegicus clone CH230-41J2, *** SEQUENCING IN PROGRESS ***,
2 unordered pieces.
ACCESSION AC133845
VERSION AC133845.3 GI:30522110
KEYWORDS HTG: HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 234325)
AUTHORS Muzny, D., Marie, Metzker, M., Lee, Abramson, S., Adams, C., Alder, J.,
Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D.,
Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,
Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
Brvant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,
Cardenas, V., Carter, K., Cavazos, I., Caesar, H., Center, A.,
Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,
Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
Davila, M., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,
Egan, A., Escoto, M., Eugene, C., Evans, C., Falls, T., Fan, G.,
Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
Fraser, C., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
Georgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W.,
Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K.,
Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,
Hernandez, R., Hines, S., Hladun, S., Hodgson, A., Hughes, M.,
Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A.,
Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,
Karpthy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,
Kowis, C., Kraft, C., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,
Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,
Lorensheuer, L., Loulseged, H., Lozado, R., Lu, X., Ma, J.,
Maheshwari, M., Mahindarne, M., Mahmoud, M., Malloy, K., Mangum, A.,
Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,
Mawhney, S., McLeod, M., McNeill, T., Meenen, E.,
Miloevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,
Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L.,
Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Parks, K.,
Nwaokeme, O., Okwuonu, G., Olarnpunsagoon, A., Pal, S., Parks, K.,
Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C.,
Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.,
Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M., Reigh, R.,
Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, P.,
Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.,
Sanders, W., Savary, G., Scherer, S., Scott, G., Shatsman, S., Shen, H.,
Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C., Smales, D.,
Speed, A., Sodergren, E., Song, X., Sorelle, R., Sosa, J.,
Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C.,
Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K.,
```

Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wlecyk, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstein, G. and Gibbs, R.A.

Direct Submission

Unpublished

2 (bases 1 to 234325)

Rat Genome Sequencing Consortium.

Rat Genome Sequencing Consortium.

Direct Submission

Submitted (19-SEP-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 234325)

Rat Genome Sequencing Consortium.

Rat Genome Sequencing Consortium.

Direct Submission

Submitted (10-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

On May 10, 2003 this sequence version replaced gi:25013050.

The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: GSVF

Center clone name: CH230-41J2

----- Summary Statistics

Assembly program: Atlas 3.0;

Consensus quality: 224086 bases at least Q40

Consensus quality: 226523 bases at least Q30

Consensus quality: 227917 bases at least Q20

Estimated insert size: 237878; sum-of-contigs estimation

Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently consists of 2 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

* 1 231823: contig of 231823 bp in length

* 231824 231923: gap of unknown length

* 231924 234325: contig of 2402 bp in length.

----- Location/Qualifiers

1..234325

/organism="Rattus norvegicus"

/mol_type="genomic DNA"

/db_xref="taxon:10116"

/clone="CH230-41J2"

1..3168

/note="wgs_end extension

clone_end:Sp6"

6731..53657

/note="clone boundary
 clone_end:Sp6
 site:ECORI
 end_sequence: BH342505"
 complement(230261..231118)
 /note="clone boundary
 clone_end:T7
 site:ECORI
 end_sequence: BH342503"

misc_feature

ORIGIN

Query Match 12.4%; Score 22; DB 2; Length 234325;
 Best Local Similarity 100.0%; Pred. No. 0.23;

Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 76 GACTGCATTAGATCTGTCACA 97

|||||

Db 54847 GACTGCATTAGATCTGTCACA 54868

RESULT 4

AC094603

LOCUS

AC094603

DEFINITION

Rattus norvegicus clone CH230-3J24, WORKING DRAFT SEQUENCE.

ACCESSION

AC094603

VERSION

AC094603.6 GI:30466316

KEYWORDS

HTG, HTGS_PHASE2, HTGS_DRAFT; HTGS_FULLTOP.

SOURCE

Rattus norvegicus

ORGANISM

Rattus norvegicus

REFERENCE

1 (bases 1 to 235305)

AUTHORS

Muzny, D. Marie., Metzker, M. Lee., Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Alebrooks, S., Amin, A., Anguiano, D., Anyalebechi, V., Ayagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C. M., Gabis, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebregeorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W., Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Haves, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hogues, M., Hollins, B., Howells, S., Hulyk, S., Hume, J., Idiebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpachy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorenshuhwa, L., Loulseghe, H., Lozado, R. J., Lu, X., Ma, J., Maheshwari, M., Mahindratne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mawhney, S., McLeod, M. P., McNeill, T. Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwaokeme, O., Okwuonu, G., Olarnpunsagoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.-L., Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J., Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C. D., Snajjs, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J.,

Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Umani, K., Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleciyk, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, P., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G. and Gibbs, R.A.

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

2 (bases 1 to 235305)
Unpublished
Worley, K.C.

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

3 (bases 1 to 235305)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (09-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

COMMENT

On May 9, 2003 this sequence version replaced gi:24818504.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GAZJ
Center clone name: CH230-3J24
----- Summary Statistics
Assembly program: Atlas;
Consensus quality: 230918 bases at least Q40
Consensus quality: 231629 bases at least Q30
Consensus quality: 232106 bases at least Q20
Estimated insert size: 240688; sum-of-contigs estimation
Quality coverage: 9x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html.
* NOTE: This is a 'working draft' sequence. It currently consists of 1 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submittor.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 235305: contig of 235305 bp in length.
* Location/Qualifiers
1. 235305
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-3J24"
complement(1308..2023)
/note="clone_boundary
clone_end:77

FEATURES
source

misc_feature

site: EcoRI
end_sequence: BH311502"
complement(233651..234141)
/note="clone boundary
clone_end:Spf6
site: EcoRI
end_sequence: BH311503"

ORIGIN

Query Match 12.4% Score 22; DB 2; Length 235305;
Best Local Similarity 100.0%; Pred. No. 0.23;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 81 CATTAGATCTGTCAGAAACAC 102
|||||
Db 143087 CATTAGATCTGTCAGAAACAC 143108
|||||

RESULT 5
AC094293/c
LOCUS
DEFINITION Rattus norvegicus clone CH230-3K24, *** SEQUENCING IN PROGRESS ***, 20 unordered pieces.
AC094293 AC094293.7 GI:30467443
HTG; HTGS PHAS1; HTGS DRAFT; HTGS_ENRICHED.
KEYWORDS Rattus norvegicus (Norway rat)
SOURCE Rattus norvegicus
ORGANISM Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

REFERENCE
1 (bases 1 to 246625)
Muzny, D., Marie, E., Metzker, M., Lee, J., Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D., Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Bryant, N., Burch, P., Burrell, K., Blyth, P., Brown, M., Cardenas, V., Carter, K., Cavazos, I., Caesar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Evans, K., Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebregeorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W., Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, J., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogues, M., Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpach, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorensuena, L., Louised, H., Lozado, R.J., Lu, X., Ma, J., Maheshwari, N., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mawhney, S., McLeod, M.P., McNeill, T.Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwaokemele, O., Okwuonu, G., Olarnpunsagoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plopper, P., Poindexter, A., Popovic, D., Primus, E., Pu, L., Puzo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojars, A., Rose, M., Rose, R., Ruiz, S.J., Sanders, W., Savary, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajs, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C.,

Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wlezyk, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, B., Yoon, V., Yu, F., Zhang, J., Zhou, X., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G. and Gibbs, R.A.

Direct Submission
2 (bases 1 to 246625)
Worley, K.C.

Direct Submission
Submitted (17-SEP-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 246625)
Rat Genome Sequencing Consortium.

Direct Submission
Submitted (09-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

On May 9, 2003 this sequence version replaced gi:25188904.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GAIN
Center clone name: CH230-3K24
----- Summary Statistics
Assembly program: Atlas;
Consensus quality: 184728 bases at least Q40
Consensus quality: 191694 bases at least Q30
Consensus quality: 196691 bases at least Q20
Estimated insert size: 187878; sum-of-contigs estimation
Quality coverage: 5x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently consists of 20 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

* 1 205043: contig of 205043 bp in length
* 205044 205143: gap of unknown length
* 205144 215060: contig of 9917 bp in length
* 215061 215160: gap of unknown length
* 215161 216255: contig of 1095 bp in length
* 216256 216355: gap of unknown length
* 216356 217394: contig of 1039 bp in length
* 217395 217494: gap of unknown length
* 217495 218555: contig of 1061 bp in length
* 218556 218656: gap of unknown length
* 218656 220239: contig of 1584 bp in length
* 220240 220339: gap of unknown length

* 220340 222257: contig of 1918 bp in length
* 222258 222357: gap of unknown length
* 222358 223852: contig of 1495 bp in length
* 223853 223952: gap of unknown length
* 223953 225460: contig of 1508 bp in length
* 225461 225560: gap of unknown length
* 225561 226912: contig of 1352 bp in length
* 226913 227012: gap of unknown length
* 227013 228059: contig of 1047 bp in length
* 228060 228159: gap of unknown length
* 228160 229456: contig of 1297 bp in length
* 229457 229556: gap of unknown length
* 229557 230918: contig of 1362 bp in length
* 230919 231018: gap of unknown length
* 231019 232608: contig of 1590 bp in length
* 232609 232708: gap of unknown length
* 232709 233944: contig of 1236 bp in length
* 233945 234044: gap of unknown length
* 234045 237241: contig of 3197 bp in length
* 237242 237341: gap of unknown length
* 237342 238817: contig of 1476 bp in length
* 238818 238917: gap of unknown length
* 238918 240337: contig of 1420 bp in length
* 240338 240437: gap of unknown length
* 240438 242451: contig of 2014 bp in length
* 242452 242551: gap of unknown length
* 242552 246625: contig of 4074 bp in length.

FEATURES
Location/Qualifiers
source
1. 246625
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-3K24"
misc_feature
113870..115506
/note="wgs_contig"
misc_feature
157302..158702
/note="wgs_contig"
misc_feature
188107..190360
/note="wgs_contig"
misc_feature
213846..215060
/note="wgs_contig"
ORIGIN
Query Match 12.4% Score 22; DB 2; Length 246625;
Best Local Similarity 100.0%; Pred. No. 0.23; Indels 0; Gaps 0;
Matches 22; Conservative 0; Mismatches 0;
Qy 81 CATTAGATCTGTCTGAGAAACAC 102
|||||
Db 65928 CATTAGATCTGTCTGAGAAACAC 65907

RESULT 6
AC117346/c
LOCUS
DEFINITION Rattus norvegicus clone CH230-391F21, WORKING DRAFT SEQUENCE, 3 linear HTG 19-NOV-2002
AC117346
unorderd pieces.
ACCESSION AC117346 GI:24635507
VERSION AC117346.6
KEYWORDS HTG; HTGS PHASE1; HTGS DRAFT; HTGS_FULLTOP.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE
1 (bases 1 to 205947)
AUTHORS
Muzny, D., Marie, Metzker, M., Lee, Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Alebrooks, S., Amin, A., Angiano, D., Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Biewald, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burrell, K., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A.,

Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,
 Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
 Davila, M.L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
 Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
 Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,
 Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G.,
 Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
 Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
 Gebregorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W.,
 Gunaratne, P., Haaland, P., Hawes, A., Hamilton, C., Hamilton, K.,
 Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,
 Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogue, M.,
 Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A.,
 Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,
 Karpach, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,
 Kowitz, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,
 Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,
 Lorenshwaha, L., Loulseghe, H., Lozado, R.J., Lu, X., Ma, J.,
 Maheshwari, M., Mahindaratne, M., Mahmoud, M., Malloy, K., Mangum, A.,
 Mangum, B., Mapua, P., Martin, K., Martin, R., Malloy, K., Mangum, A.,
 Mawhinney, S., McLeod, M.P., McNeill, T.Z., Meenen, E.,
 Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,
 Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L.,
 Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,
 Nwankwelu, O., Okwuonu, G., Olarnpungoon, A., Pal, S., Parks, K.,
 Paeternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C.,
 Plummer, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.-L.,
 Puzo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R.,
 Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,
 Rivera, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J.,
 Sanders, W., Savary, G., Scherer, S., Scott, G., Shateman, S., Shen, H.,
 Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajs, D.,
 Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J.,
 Steimle, M., Strong, R., Sutton, A., Svatek, A., Taber, P., Taylor, C.,
 Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K.,
 Valas, R., Vera, V., Villalana, D., Waldron, L., Walker, B., Wang, J.,
 Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,
 Williams, G., Willson, R., Wlezyk, R., Wooden, H., Worley, K.,
 Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,
 Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von
 Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O.,
 Weinstock, G. and Gibbs, R.A.
 Direct Submission
 Unpublished
 2 (bases 1 to 205947)
 Worley, K.C.
 Direct Submission
 Submitted (10-APR-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 205947)
 Rat Genome Sequencing Consortium.
 Direct Submission
 Submitted (19-NOV-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 On Nov 6, 2002 this sequence version replaced gi:23812771.
 The sequence in this assembly is a combination of BAC based reads
 and whole genome shotgun sequencing reads assembled using Atlas
 (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
 in the feature table below represents a scaffold in the Atlas
 assembly (a 'contig-scaffold'). Within each contig-scaffold,
 individual sequence contigs are ordered and oriented, and separated
 by sized gaps filled with Ns to the estimated size. The sequence
 may extend beyond the ends of the clone and there may be sequence
 contigs within a contig-scaffold that consist entirely of whole
 genome shotgun sequence reads. Both end sequences and whole genome
 shotgun sequence only contigs will be indicated in the feature
 table.

----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: http://www.hgsc.bcm.tmc.edu/

Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 Center project name: GUSD
 Center clone name: CH230-391F21
 ----- Summary Statistics
 Assembly program: Phrap; version 0.990329
 Consensus quality: 197835 bases at least Q40
 Consensus quality: 199693 bases at least Q30
 Consensus quality: 200655 bases at least Q20
 Estimated insert size: 202041; sum-of-contigs estimation
 Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

 * NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 3 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

FEATURES	source	1..205947
	organism="Rattus norvegicus"	
	/mol_type="genomic DNA"	
	/db_xref="taxon:10116"	
	/clone="CH230-391F21"	
	1..1121	
misc_feature	/notes="wgs_end_extension"	
	clone_end:T7"	
	1478..-2516	
misc_feature	/note="wgs_end_extension"	
	clone_end:T7"	
	2275..-72955	
misc_feature	/notes="clone boundary"	
	clone_end:T7"	
	site:	
	end sequence:BZ2111748"	
misc_feature	complement(5314..-6288)	
	/notes="clone boundary"	
	clone_end:Sp6	
	site:	
	end sequence:BZ2111750"	
misc_feature	85851..88132	
	/notes="wgs_end_extension"	
	clone_end:Sp6"	
ORIGIN		
Query Match	11.9%;	Score 21; DB 2; Length 205947;
Best Local Similarity	100.0%;	Pred. No. 0.87;
Matches	21; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
Qy	144 TGAATTTCTGAAGGATGAGAA 164	
Db	64420 TGAATTTCTGAAGGATGAGAA 64400	
RESULT 7		
AC114722/c		
LOCUS	AC114722	291547 bp DNA linear HTG 19-NOV-2002
DEFINITION	Rattus norvegicus clone CH230-60K20, WORKING DRAFT SEQUENCE.	
ACCESSION	AC114722	
VERSION	AC114722.4 GI:25072902	
KEYWORDS	HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_FULLTOP.	
SOURCE	Rattus norvegicus (Norway rat)	
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

1 (bases 1 to 291547)
Muzny D, Marie, Metzker, M. Lee., Abramson, S., Adams, C., Alder, J.,
Allen, C., Allen, H., Albrooks, S., Amin, A., Anguiano, D.,
Ayalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,
Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,
Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A.,
Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,
Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,
Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G.,
Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, P.,
Fraser, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
Gebregregis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W.,
Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K.,
Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,
Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hogues, M.,
Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A.,
Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,
Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,
Kowis, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,
Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,
Lorensuewa, L., Loulseged, H., Lozano, R. J., Lu, X., Ma, J.,
Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A.,
Mangum, B., Mapa, P., Martin, K., Martin, R., Martinez, E.,
Mawhinney, S., McLeod, M. F., McNeill, T. Z., Meenen, E.,
Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,
Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L.,
Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Parks, K.,
Nwackeleme, O., Okwuonu, G., Olarnpunsagoon, A., Pal, S., Parks, K.,
Pasternak, S., Paul, H., Perez, A., Perez, L., Pfamkoch, C.,
Plopper, F., Poindecker, A., Popovic, D., Primus, E., Pu, L.-L.,
Puzo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R.,
Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,
Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J.,
Sanders, M., Savary, G., Scherer, S., Scott, G., Shatsman, S., Shen, H.,
Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C. D., Smajs, D.,
Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J.,
Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C.,
Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K.,
Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J.,
Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,
Williams, G., Willson, R., Wleczky, R., Wooden, H., Worley, K.,
Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,
Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Zhao, S., Dunn, D., von
Niederhausen, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O.,
Weinstock, G. and Gibbs, R. A.

TITLE

JOURNAL

REFERENCE

Unpublished

2 (bases 1 to 291547)

Worley, K. C.

TITLE

JOURNAL

Submitted (11-MAR-2002)

Human Genome Sequencing Center, Department

of Molecular and Human Genetics, Baylor College of Medicine, One

Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 291547)

Rat Genome Sequencing Consortium.

REFERENCE

Unpublished

2 (bases 1 to 291547)

Worley, K. C.

TITLE

JOURNAL

Submitted (19-NOV-2002)

Human Genome Sequencing Center, Department

of Molecular and Human Genetics, Baylor College of Medicine, One

Baylor Plaza, Houston, TX 77030, USA

On Nov 19, 2002 this sequence version replaced gi:23265675.

by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GTJR
Center clone name: CH230-60K20
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 231275 bases at least Q40
Consensus quality: 234311 bases at least Q30
Consensus quality: 235939 bases at least Q20
Estimated insert size: 240124; sum-of-contigs estimation
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.

* 1 291547: contig of 291547 bp in length.

FEATURES

source

1. .291547

/organism="Rattus norvegicus"

/mol_type="genomic DNA"

/db_xref="taxon:10116"

/clone="CH230-60K20"

1. .1044

/notes="wgs end_extension

clone_end:Sp6"

complement(8119..8720)

/notes="clone boundary

clone_end:Sp6

site:

end sequence:BH365931"

21921..119834

/notes="clone boundary

clone_end:T7

site:

end sequence:BH365929"

32443..34180

/notes="wgs end_extension

clone_end:T7"

misc_feature

Query Match 11.9%; Score 21; DB 2; Length 291547;

Best Local Similarity 100.0%; Pred. No. 0.85;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

ORIGIN

144 TGAATTTCTGAAGTGAAGAA 164

13561 TGAATTTCTGAAGTGAAGAA 13541

RESULT 8

AC015339/c

LOCUS

AC015339 45550 bp DNA linear HTG 16-NOV-1999

DEFINITION Drosophila melanogaster, *** SEQUENCING IN PROGRESS ***.

AC015339

AC015339.1 GI:6435996

KEYWORDS HTG; HTGS_PHASE2.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 45550)
AUTHORS Adams,M. and Venter,J.C.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-1999) Celera Genomics, 45 West Gude Drive, Rockville, MD, USA
COMMENT This sequence was identified as CDM:10209553 by the submitter.
* NOTE: This is a "working draft" sequence.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.

FEATURES
source
1..45550
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"

ORIGIN
Query Match 11.3%; Score 20; DB 2; Length 45550;
Best Local Similarity 100.0%; Pred. No. 3.7;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 127 GCCTTCCCAATGCTGGTGTGA 146
|||||
Db 34535 GCCTTCCCAATGCTGGTGTGA 34516
|||||

RESULT 9
AC008337/c
LOCUS 169680 bp DNA linear INV 31-JUL-2004
DEFINITION Drosophila melanogaster clone BACR02B03, complete sequence.
ACCESSION AC008337
VERSION AC008337.8 GI:50872396
KEYWORDS HTG.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 169680)
AUTHORS Celniker,S.E., Agbayani,A., Arcaina,T.T., Baxter,E., Blazej,R.G., Butenhoff,C., Champe,M., Chavez,C., Chew,M., Ciesiolka,L., Doyle,C.M., Farfan,D.E., Galle,R., George,R.A., Harris,N.L., Hoskins,R.A., Houston,K.A., Hummasti,S.R., Karra,K., Kearney,L., Kim,E., Lee,B., Lewis,S., Li,P., Lomotan,M.A., Mazda,P., Moshrefi,A.R., Moshrefi,M., Nixon,K., Pacleb,J.M., Park,S., Pfeiffer,B., Poon,L., Sequeira,A., Sethi,H., Snir,E., Svirkas,R.R., Wan,K.H., Weinburg,T., Zhang,R., Zieran,L.L. and Rubin,G.M.
TITLE Direct Submission
JOURNAL Submitted (02-AUG-1999) Drosophila Genome Center, Lawrence Berkeley Laboratory, MS 64-121, Berkeley, CA 94720, USA
REFERENCE 2 (bases 1 to 169680)
AUTHORS Celniker,S., Carlson,J., Wan,K., Pfeiffer,B., Frise,E., George,R., Hoskins,R., Stapleton,M., Pacleb,J., Park,S., Svirkas,R., Smith,E., Yu,C. and Rubin,G.
TITLE Direct Submission
JOURNAL Submitted (31-JUL-2004) Berkeley Drosophila Genome Project, MS 64-121, Lawrence Berkeley National Laboratory, One Cyclotron Road, Berkeley, CA 94720, US
COMMENT On Jul 31, 2004 this sequence version replaced gi:15451507.
Sequence submitted by:
Berkeley Drosophila Genome Project
Lawrence Berkeley National Laboratory
Berkeley, CA 94720
This sequence submission incorporates changes made during the reevaluation of the assembly or fingerprint verification of the

clone. For further information about this sequence, including its location and relationship to other sequences, please visit our sequence archive Web site (<http://www.fruitfly.org/sequence/>) or send email to bdgp@fruitfly.org.

FEATURES
source
1..169680
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/strain="y; cn bw sp"
/db_xref="taxon:7227"
/chromosome="X"
/map="18A-18B"
/clone="BAC clone BACR02B03 (D902)"
/clone_lib="RPCI-98 (Roswell Park Cancer Institute Drosophila melanogaster BAC library, partial EcoRI in pBACE3.6)"

ORIGIN
Query Match 11.3%; Score 20; DB 3; Length 169680;
Best Local Similarity 100.0%; Pred. No. 3.4;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 127 GCCTTCCCAATGCTGGTGTGA 146
|||||
Db 113927 GCCTTCCCAATGCTGGTGTGA 113908
|||||

RESULT 10
AC012376/c
LOCUS 183048 bp DNA linear INV 31-JUL-2004
DEFINITION Drosophila melanogaster clone BACR48C12, complete sequence.
ACCESSION AC012376
VERSION AC012376.12 GI:50872344
KEYWORDS HTG.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 183048)
AUTHORS Celniker,S.E., Agbayani,A., Arcaina,T.T., Baxter,E., Blazej,R.G., Butenhoff,C., Champe,M., Chavez,C., Chew,M., Ciesiolka,L., Doyle,C.M., Farfan,D.E., Galle,R., George,R.A., Harris,N.L., Hoskins,R.A., Houston,K.A., Hummasti,S.R., Karra,K., Kearney,L., Kim,E., Lee,B., Lewis,S., Li,P., Lomotan,M.A., Mazda,P., Moshrefi,A.R., Moshrefi,M., Nixon,K., Pacleb,J.M., Park,S., Pfeiffer,B., Poon,L., Sequeira,A., Sethi,H., Snir,E., Svirkas,R.R., Wan,K.H., Weinburg,T., Zhang,R., Zieran,L.L. and Rubin,G.M.
TITLE Direct Submission
JOURNAL Submitted (26-OCT-1999) Drosophila Genome Center, Lawrence Berkeley Laboratory, MS 64-121, Berkeley, CA 94720, USA
REFERENCE 2 (bases 1 to 183048)
AUTHORS Celniker,S., Carlson,J., Wan,K., Pfeiffer,B., Frise,E., George,R., Hoskins,R., Stapleton,M., Pacleb,J., Park,S., Svirkas,R., Smith,E., Yu,C. and Rubin,G.
TITLE Direct Submission
JOURNAL Submitted (31-JUL-2004) Berkeley Drosophila Genome Project, MS 64-121, Lawrence Berkeley National Laboratory, One Cyclotron Road, Berkeley, CA 94720, US
COMMENT On Jul 31, 2004 this sequence version replaced gi:13324739.
Sequence submitted by:
Berkeley Drosophila Genome Project
Lawrence Berkeley National Laboratory
Berkeley, CA 94720
This sequence submission incorporates changes made during the reevaluation of the assembly or fingerprint verification of the clone. For further information about this sequence, including its location and relationship to other sequences, please visit our sequence archive Web site (<http://www.fruitfly.org/sequence/>) or send email to bdgp@fruitfly.org.

FEATURES
source
1..183048
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/strain="y; cn bw sp"
/db_xref="taxon:7227"
/chromosome="X"
/map="18A-18B"
/clone="BAC clone BACR02B03 (D902)"
/clone_lib="RPCI-98 (Roswell Park Cancer Institute Drosophila melanogaster BAC library, partial EcoRI in pBACE3.6)"

```
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/strain="Y: cn bw sp"
/db_xref="taxon:7227"
/chromosome="X"
/map="18B-18B"
/clones="BAC clone BACR48C12 (D1122)"
/clone_lib="RPCI-98 (Roswell Park Cancer Institute
Drosophila melanogaster BAC library, partial EcoRI in
pBACE3.6)"
```

ORIGIN

```
Query Match      : 11.3%; Score 20; DB 3; Length 183048;
Best Local Similarity 100.0%; Pred. No. 3.3;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 127 GCCTTCCCAATGCTGGTGTGA 146
      |||||
Db 66140 GCCTTCCCAATGCTGGTGTGA 66121
```

RESULT 11

```
AC119359          207871 bp  DNA  linear  HTG 20-NOV-2002
LOCUS             Rattus norvegicus clone CH230-466N16, WORKING DRAFT SEQUENCE, 2
DEFINITION        unordered pieces.
```

```
ACCESSION          AC119359
VERSION            AC119359.5 GI:25137741
```

```
KEYWORDS           HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
```

```
SOURCE             Rattus norvegicus (Norway rat)
```

```
ORGANISM            Rattus norvegicus
```

```
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
```

```
1 (bases 1 to 207871)
```

REFERENCE

AUTHORS

```
Munzy,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J.,
Allen,C., Allen,H., Albrooks,S., Amin,A., Anguiano,D.,
Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,P.,
Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
Chacko,J., Chavez,D., Chen,G., Chen,X., Chen,Z., Chu,J.,
Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
Delgado,O., Dengon,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
Gebrtegeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
Gunatratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
Karpachy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
Lorenshwari,L., Loulseghe,H., Lozado,R.J., Lu,X., Ma,J.,
Maheshwari,M., Mahindaratne,M., Mahmoud,M., Malloy,K., Mangum,A.,
Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
Mawhinney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
Nwaokemele,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K.,
Pasternak,S., Paul,H., Perez,A., Perez,I., Pfannkoch,C.,
Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L.,
Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
Rivers,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
```

```
Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smaje,D.,
Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,
Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K.,
Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
Yu,F., Zhang,J., Zhou,X., Zhao,S., Dunn,D., von
Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
Weinstock,G. and Gibbs,R.A.
```

Direct Submission

Unpublished

```
2 (bases 1 to 207871)
```

AUTHORS

```
Worley,K.C.
```

Direct Submission

```
Submitted (26-APR-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
```

```
3 (bases 1 to 207871)
```

Rat Genome Sequencing Consortium.

Direct Submission

```
Submitted (20-NOV-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
```

```
On Nov 20, 2002 this sequence version replaced gi:23908428.
```

The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

```
----- Genome Center of Medicine
```

```
Center: Baylor College of Medicine
```

```
Center code: BCM
```

```
Web site: http://www.hgsc.bcm.tmc.edu/
```

```
Contact: hgsc-help@bcm.tmc.edu
```

```
----- Project Information
```

```
Center project name: GVFX
```

```
Center clone name: CH230-466N16
```

```
----- Summary Statistics
```

```
Assembly program: Phrap; version 0.990329
```

```
Consensus quality: 177278 bases at least Q40
```

```
Consensus quality: 180313 bases at least Q30
```

```
Consensus quality: 181910 bases at least Q20
```

```
Estimated insert size: 183823; sum-of-contigs estimation
```

```
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation
```

```
-----
```

```
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank\_draft\_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 2 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
```

```
* 1 206612: contig of 206612 bp in length
```

```
* 206613 206712: gap of unknown length
```

```
* 206713 207871: contig of 1159 bp in length.
```

FEATURES

source

```
1..207871
    Location/Qualifiers
        /organism="Rattus norvegicus"
        /mol_type="genomic DNA"
        /db_xref="taxon:10116"
        /clone="CH230-466N16"
```

```

misc_feature      1. .1047
                  /note="wgs_end_extension
clone_end:Sp6"
misc_feature      24305. .25227
                  /note="clone_boundary
clone_end:Sp6
site:
end_sequence:BZ216496"
misc_feature      45252. .46059
                  /note="clone_boundary
clone_end:T7
site:
end_sequence:BZ216494"
misc_feature      205108. .206612
                  /note="wgs_end_extension
clone_end:T7"

ORIGIN
Query Match      11.3%; Score 20; DB 2; Length 207871;
Beet Local Similarity 100.0%; Pred. No. 3.3;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 141 GTGTGAATTTCTGAAGGATG 160
      |||||
Db 118039 GTGTGAATTTCTGAAGGATG 118058

RESULT 12
AE003511/c
LOCUS
DEFINITION Drosophila melanogaster chromosome X, section 63 of 74 of the
complete sequence.
ACCESSION AE003511 AE002593 AE014298
VERSION AE003511.3 Gi:22832537
KEYWORDS
SOURCE
ORGANISM Drosophila melanogaster (fruit fly)
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 297970)
Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
Ananâtides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
Brandon R.C., Rogers Y.H., Blazej R.G., Champe M., Pfeiffer B.D.,
Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
Abriel J.F., Agbayani A., An H.J., Andrews-Pfankuch C., Baldwin D.,
Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A.,
Chandra I., Cherry J.M., Cawley S., Dahlke C., Davenport L.B.,
Davies P., de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I.,
Dietz S.M., Dodson K., Doup L.E., Downes M., Dugan-Rocha S.,
Dunkov B.C., Dunn P., Durbin K.J., Evangelista C., Ferraz C.,
Ferriera S., Fleischmann W., Fosler C., Gabrielian A.E., Garg N.S.,
Gelbart W.M., Glasser K., Glodek A., Gong F., Gorrell J.H., Gu Z.,
Guan P., Harris M., Harris N.L., Harvey D., Heiman T.J.,
Hernandez J.R., Houck J., Hostin D., Houston K.A., Howland T.J.,
Wei M.H., Ibegwam C., Jalali M., Kalush F., Karpen G.H., Ke Z.,
Kennison J.A., Ketchum K.A., Kimmel B.E., Kodira C.D., Kraft C.,
Kravitz S., Kulp D., Lai Z., Lasoko P., Lei Y., Levitsky A.A.,
Li J., Li Z., Liang Y., Lin X., Liu X., Mattei B., McIntosh T.C.,
McLeod M.P., McPherson D., Merkulov G., Milshina N.V., Moberg C.,
Morris J., Moshrefi A., Mount S.M., Moy M., Murphy B., Murphy L.,
Muzny D.M., Nelson D.L., Nelson D.R., Nelson K.A., Nixon K.,
Nusskern D.P., Pacleib J.M., Palazzolo M., Pitman G.S., Pan S.,
Pollard J., Puri V., Reese M.G., Reinert K., Remington K.,
Saunders R.D., Scheeler F., Shen H., Shue B.C., Siden-Kiamos I.,
Simpson M., Skupski M.P., Smith T., Spier E., Spradling A.C.,
Stapleton M., Strong R., Sun E., Svirskas R., Tector C., Turner R.,
Venter E., Wang A.H., Wang X., Wang Z.Y., Wasearman D.A.,
Weinstock G.M., Weissenbach J., Williams S.M., Woodage T.,
Worley K.C., Wu D., Yang S., Yao Q.A., Ye J., Yeh R.F.,
Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L., Zheng X.H.,
Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
Gibbs R.A., Myers E.W., Rubin G.M. and Venter J.C.
The genome sequence of Drosophila melanogaster
Science 287 (5461), 2185-2195 (2000)
20196006
10731132
2 (bases 1 to 297970)
Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W.,
Halpern A., Patel S., Adams M., Champe M., Dugan S.P., Frisoe E.,
Hodgson A., George R.A., Hoskins R.A., Laverty T., Muzny D.M.,
Nelson C.R., Pacleib J.M., Park S., Pfeiffer B.D., Richards S.,
Sodergren E.J., Svirskas R., Tabor P.E., Wan K., Stapleton M.,
Sutton G.G., Venter C., Weinstock G., Scherer S.E., Myers E.W.,
Gibbs R.A. and Rubin G.M.
Finishing a whole-genome shotgun: release 3 of the Drosophila
melanogaster euchromatic genome sequence
Genome Biol. 3 (12), RESEARCH0079 (2002)
22426065
12537568
3 (bases 1 to 297970)
Misra S., Crosby M.A., Mungall C.J., Matthews B., Campbell K.S.,
Hradecky P., Huang Y., Kaminker J.S., Milburn G.H., Prochnik S.E.,
Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L.,
Berman B.P., Bettencourt B.R., Celniker S.E., de Grey A.D.,
Drysdale R.A., Harris N.L., Richter J., Russo S., Schroeder A.J.,
Shu S.Q., Stapleton M., Yamada C., Ashburner M., Gelbart W.M.,
Rubin G.M. and Lewis S.E.
Annotation of the Drosophila melanogaster euchromatic genome: a
systematic review
Genome Biol. 3 (12), RESEARCH0083 (2002)
22426069
12537572
4 (bases 1 to 297970)
Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J.,
Svirskas R., Patel S., Frise E., Wheeler D.A., Lewis S.E.,
Rubin G.M., Ashburner M. and Celniker S.E.
The transposable elements of the Drosophila melanogaster
euchromatin: a genomics perspective
Genome Biol. 3 (12), RESEARCH0084 (2002)
22426070
12537573
5 (bases 1 to 297970)
Adams M.D., Celniker S.E., Gibbs R.A., Rubin G.M. and Venter C.J.
Direct Submission
Submitted (21-MAR-2000) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
6 (bases 1 to 297970)
FlyBase
Direct Submission
Submitted (06-SEP-2002) University of California Berkeley, 539 Life
Sciences Addition, Berkeley, CA 94720, USA
7 (bases 1 to 297970)
FlyBase
Direct Submission
Submitted (10-MAR-2004) FlyBase, Harvard University, Biological
Laboratories, 16 Divinity Avenue, Cambridge, MA 02138, USA
On Sep 13, 2002 this sequence version replaced gi:10728334.
FEATURES
source
1. .297970
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/chromosomes="X"
complement (5294..10011)
/locus_tag="TE19084"
/map="18A3-18A3"
/transposon="Doc{}163"
/db_xref="FLYBASE:FBti0019084"
repeat_region

```

gene	<div><10239. .>12245 /locus_tag="CG14190" /map="18A3-18A3" /db_xref="FLYBASE:FBgn0030979" 10239. .>12245 /locus_tag="CG14190" /product="CG14190-PA" /db_xref="FLYBASE:FBgn0030979" 10239. .12245 /locus_tag="CG14190" /codon_start=1 /product="CG14190-PA" /protein_id="AAF48911.1" /db_xref="GI:7293538" /db_xref="FLYBASE:FBgn0030979" /translation="MSSEMEITPTTGMDEVSEKTAERKVFAPMPEAEWSEKMDITPVVG MDEVSEKSAEKVAPMEAESEMEITPVTCMDDEVSEYEEKGIAPFKAHMSEDE NTPVTGMEVSQTYEENAIAPMSEICASGDPVSPMGVAQENETASIPDGIITRLKDI LATQTLFLVQRNETVVRNEMSQMISPLVEISQENQINDPSVNEPNQIMDTSTFTSN DASTNNVYVEQESTVVMVEQNTIIIELYDYVPPASLNNTIRPFPIPVMPYPS SPPPAPKATKGRKPGPKPRATKRRPPTPLPVNPNFLAHVNLPLENEY PKVIVQVQPVLRSDVAPKPPKVAVKEAPIPVSRPVTPNPPSRRLVLSAHLRE KLRLFGDVTISLIDRVSPEKAVREPDGSOQILFTAPQLLSFTAHRGNWQWV PIQMLMKPSAPMLSHVFPQMGQTLVKMFAQQDLKPSLPFPKPIGRPKPSKS TTPRSKTYLHPGKPVQSKFAPPAAPPVAPPAATFAPFPDPTSHPGADPAPEST PTPAPENPPDQPTLTPTAPPAVPTAPTVAASPTDPDESMPFPPIILPAIKES SQERLRKSIWAKRIKAMKEAAAAAASAKLSLNNDMPAKFP" complement (13085. .>13737) /locus_tag="CG33254" /map="18A3-18A3" /db_xref="FLYBASE:FBgn0053254" complement (13085. .>13737) /locus_tag="CG33254-RA" /product="CG33254-RA" /db_xref="FLYBASE:FBgn0053254" complement (13390. .13737) /locus_tag="CG33254" /codon_start=1 /product="CG33254-PA" /protein_id="AAS65407.1" /db_xref="GI:45447056" /db_xref="FLYBASE:FBgn0053254" /translation="MQGVPAHSAKSTRVHFASARRIPVTRSSVQSIVLRPDPFOE SKQHVVENCHRSRDWRPVLQRGQGLRLGSLAAGHAQCKGLPSETGVRIPOG GGFLKWTPTGVD" 22550. 23419 /locus_tag="CG7406" /map="18A3-18A3" /db_xref="FLYBASE:FBgn0030980" join (22550. .22636,22698. .23419) /locus_tag="CG7406" /product="CG7406-RA" /db_xref="FLYBASE:FBgn0030980" join (22631. .22636,22698. .23198) /locus_tag="CG7406" /codon_start=1 /product="CG7406-PA" /protein_id="AAF48913.1" /db_xref="GI:7293540" /db_xref="FLYBASE:FBgn0030980" /translation="MKVFICLIACFQAQSGFTGGASGKWTGGGGWSSGGGGAPT IKVVISRAGHGGWAGGYSGYAHAPVEIKVVISRAGSHGHDYHSHGSDVKI IKVIOERGHSHGHGHDYSSGGHGAHAEDVKI IKLISGGIADGSHGWSGGGM FSGSGGN" 26492. .27398 /locus_tag="CG14191" /map="18A3-18A3" /db_xref="FLYBASE:FBgn0030981" join (26492. .26574,26690. .27398) /locus_tag="CG14191" /product="CG14191-RA" /db_xref="FLYBASE:FBgn0030981" join (26569. .26574,26690. .27265)</div>
gene	<div><10239. .>12245 /locus_tag="CG14191" /codon_start=1 /product="CG14191-PA" /protein_id="AAF48914.2" /db_xref="GI:22832539" /db_xref="FLYBASE:FBgn0030981" /translation="MKVFYCLAVCSMAHAGFLGGGAGGSALSSGWSGGGGGGYSG GYSGHGGGGGGYGGGGYGGGGGGGKTIKIVTDSGAGGGGGYGGGGYGGH GGYGGGGGGGGYSGGGGGWSGGGGYGGGGYGGGKIIKIVISDAGSGGGYGGG YGGHGGGGYSSGGSYSAGGWAPOGGWAQSSN" Query Match 11.3%; Score 20; DB 3; Length 297970; Best Local Similarity 100.0%; Pred. No. 3.2; Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0; Qy 127 GCCTTCCAATGCTGGTGTGA 146 Db 269250 GCCTTCCAATGCTGGTGTGA 269231 RESULT 13 AC009558 184568 bp DNA linear PRI 31-MAR-2001 LOCUS Homo sapiens chromosome 15, clone RP11-142J21, complete sequence. DEFINITION AC009558 ACCESSION AC009558 GI:13194975 VERSION HTG. KEYWORDS Homo sapiens (human) SOURCE Homo sapiens ORGANISM Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. REFERENCE 1 (bases 1 to 184568) AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E. TITLE Homo sapiens chromosome 15, clone RP11-142J21 JOURNAL Unpublished REFERENCE 2 (bases 1 to 184568) AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M., Baker,J., Baldwin,J., Barna,N., Beckerly,R., Benn,J., Brown,A., Castle,A., Cerny,J., Colangelo,M., Collins,S., Collymore,A., Cooke,P., DeArelano,K., Depayre,E., Devon,K., Dewar,K., Donelan,L., Doyle,M., Ferreira,P., FitzHugh,W., Forrest,C., Funke,K., Gage,D., Galagan,J., Gardyna,S., Gilbert,D., Grant,G., Hagos,B., Headford,A., Horton,L., Howland,J.C., Jones,C., Kann,L., Karatas,A., Lehoczy,J., Lieu,C., Locke,K., Macdonald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrim,J., Molla,M., Morris,W., Morrow,J., Mychaleckyj,J., Naylor,J., Niloff,M., O'Connor,T., O'Donnell,P., Pavlin,B., Peterson,K., Pollara,V., Riley,R., Roberts,D., Roy,A., Severy,P., Stange-Thomann,N., Stojanovic,N., Stone,C., Subramanian,A., Tesfaye,S., Torruella-Miller,I., Vassiliev,H., Vo,A., Wagner,A., Wheeler,J., Wu,X., Wyman,D., Ye,W.J. and Zody,M. Direct Submission Submitted (27-AUG-1999) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA REFERENCE 3 (bases 1 to 184568) AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S., Barna,N., Bastien,V., Boguslavsky,L., Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J., Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L., Grand-Pierre,N., Hagos,B., Headford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Karatas,A., LaRocque,K., Lamarez,R., Landers,T., Lehoczy,J., Levine,R., Liu,G., MacLean,C., Macdonald,P., Marquis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupbach,R., Seaman,S., Severy,P., Sougnuez,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S.,</div>

	Theodore,J., Travers,M., Travis,N., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
TITLE	Direct Submission
JOURNAL	Submitted (31-MAR-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT	On Mar 4, 2001 this sequence version replaced gi:13176254. All repeats were identified using RepeatMasker: Smit, A.P.A. & Green, P. (1996-1997) http://ftp.genome.washington.edu/RM/RepeatMasker.html ----- Genome Center Center: Whitehead Institute/ MIT Center for Genome Research Center code: WIBR Web site: http://www-seq.wi.mit.edu Contact: sequence.submissions@genome.wi.mit.edu ----- Project Information Center project name: L1678 Center clone name: 142_J_21 -----
FEATURES	Location/Qualifiers
source	1. 184568 /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606" /chromosome="15" /map="15" /clone="RP11-142J21" /clone_lib="RPC1-11 Human Male BAC" repeat_region 1095..1130 /rpt_family="(TAA)n" complement(1133..1353) repeat_region /rpt_family="MIR" complement(11469..1590) repeat_region /rpt_family="MER45A" complement(1791..2156) repeat_region /rpt_family="THE1C" 3640..3936 /rpt_family="AluSx" 3917..4023 /rpt_family="GA-rich" 4299..4424 /rpt_family="MER91C" complement(4597..4895) repeat_region /rpt_family="L2" 5032..5245 /rpt_family="MIR" 5586..5606 /rpt_family="AT-rich" 5715..5769 /rpt_family="(CA)n" 5771..5796 /rpt_family="(GA)n" complement(6505..6675) repeat_region /rpt_family="AluYa5" 7057..7222 /rpt_family="L1ME3" complement(7943..8117) repeat_region /rpt_family="MIR" 8287..8573 /rpt_family="AluSx" 8606..8922 /rpt_family="AluSx" complement(9851..10096) repeat_region /rpt_family="MIR" 16256..16556 /rpt_family="AluSx" complement(17636..17701) repeat_region /rpt_family="L2" 17768..17799 /rpt_family="AT-rich" complement(17807..18096) repeat_region /rpt_family="AluSp" 20270..20562


```
misc_feature      12119..15745
/note="assembly_fragment"
misc_feature      15846..19753
/note="assembly_fragment"
misc_feature      19854..28155
/note="assembly_fragment"
misc_feature      28256..36319
/note="assembly_fragment"
clone_end:17
vector_side:left"
misc_feature      36420..47349
/note="assembly_fragment"
misc_feature      47450..61276
/note="assembly_fragment"
misc_feature      61377..80216
/note="assembly_fragment"
misc_feature      80317..101592
/note="assembly_fragment"
misc_feature      101693..125692
/note="assembly_fragment"
misc_feature      125793..151472
/note="assembly_fragment"
misc_feature      151573..186716
/note="assembly_fragment"

ORIGIN
Query Match      10.7%; Score 19; DB 2; Length 186716;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      142  TGTGAATTTCTGAAGGATG 160
|||||
DB      131167  TGTGAATTTCTGAAGGATG 131185
```

```
RESULT 15
AC150648/c
LOCUS      AC150648      190414 bp      DNA      linear      ROD 15-SEP-2004
DEFINITION Mus musculus chromosome 7 clone RP23-235N5, complete sequence.
ACCESSION  AC150648
VERSION    AC150648.2  GI:52077975
KEYWORDS   HTG.
SOURCE     Mus musculus (house mouse)
ORGANISM   Mus musculus
REFERENCE  1  (bases 1 to 190414)
AUTHORS    Wilson,R.K.
TITLE      The sequence of Mus musculus clone
JOURNAL    Unpublished
REFERENCE  2  (bases 1 to 190414)
AUTHORS    Wilson,R.K.
TITLE      Direct Submission
JOURNAL    Submitted (29-JUL-2004) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
REFERENCE  3  (bases 1 to 190414)
AUTHORS    Wilson,R.K.
TITLE      Direct Submission
JOURNAL    Submitted (15-SEP-2004) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
COMMENT    On Sep 15, 2004 this sequence version replaced gi:50811813.
```

```
----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu
Contact: submissions@wustl.edu
----- Project Information -----
Center project name: M_BA0235N05
-----
FEATURES             Location/Qualifiers
     source            1..190414
                        /organism="Mus musculus"
```

```
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="7"
/clone="RP23-235N5"

ORIGIN
Query Match      10.7%; Score 19; DB 10; Length 190414;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      127  GCCTTCCAATGCTGGTGTG 145
|||||
DB      106866  GCCTTCCAATGCTGGTGTG 106848

Search completed: August 26, 2005, 22:22:33
Job time : 1957.21 secs
```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 26, 2005, 18:32:17 ; Search time 414.026 Seconds
(without alignments)
2797.232 Million cell updates/sec

Title: US-10-617-978-14_COPY_64_240

Perfect score: 177

Sequence: 1 gctgacgtccggaaacta.....atgagaacgtgaaggtctga 177

Scoring table: OLIGO NUC

Gapop 60.0 , Gapext 60.0

Searched: 7331713 seqs, 327154945 residues

Word size : 0

Total number of hits satisfying chosen parameters: 14663426

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Published Applications NA:*

1: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
2: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq.*
17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq.*
18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq.*
19: /cgn2_6/ptodata/2/pubpna/US10G_PUBCOMB.seq.*
20: /cgn2_6/ptodata/2/pubpna/US10H_PUBCOMB.seq.*
21: /cgn2_6/ptodata/2/pubpna/US10I_PUBCOMB.seq.*
22: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
23: /cgn2_6/ptodata/2/pubpna/US11A_PUBCOMB.seq.*
24: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
25: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
26: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	19	10.7	665	13	US-10-617-978-14_COPY_64_240
2	19	10.7	665	17	US-10-617-978-14_COPY_64_240
3	19	10.7	143412	13	US-10-617-978-14_COPY_64_240
4	18	10.2	195	21	US-10-617-978-14_COPY_64_240
5	18	10.2	195	21	US-10-617-978-14_COPY_64_240
6	18	10.2	195	21	US-10-617-978-14_COPY_64_240
7	18	10.2	195	21	US-10-617-978-14_COPY_64_240

RESULT 1

US-10-617-978-14_COPY_64_240
; Sequence 210235, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR FILING DATE: 2000-07-12
; PRIOR FILING DATE: 2000-07-12
; PRIOR FILING DATE: 2000-04-20
; PRIOR FILING DATE: 2000-03-29
; PRIOR FILING DATE: 2000-03-29
; PRIOR FILING DATE: 2000-02-24
; PRIOR FILING DATE: 1999-11-23
; PRIOR FILING DATE: 1999-09-28
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 210235
; LENGTH: 665
; TYPE: DNA

ALIGNMENTS

8	18	10.2	195	21	US-10-617-978-14_COPY_64_240	Sequence 171, App
9	18	10.2	198	21	US-10-617-978-14_COPY_64_240	Sequence 155, App
10	18	10.2	198	21	US-10-617-978-14_COPY_64_240	Sequence 155, App
11	18	10.2	198	21	US-10-617-978-14_COPY_64_240	Sequence 163, App
12	18	10.2	272	18	US-10-617-978-14_COPY_64_240	Sequence 13962, A
13	18	10.2	319	21	US-10-617-978-14_COPY_64_240	Sequence 153, App
14	18	10.2	323	21	US-10-617-978-14_COPY_64_240	Sequence 137, App
15	18	10.2	323	21	US-10-617-978-14_COPY_64_240	Sequence 141, App
16	18	10.2	323	21	US-10-617-978-14_COPY_64_240	Sequence 145, App
17	18	10.2	323	21	US-10-617-978-14_COPY_64_240	Sequence 157, App
18	18	10.2	323	21	US-10-617-978-14_COPY_64_240	Sequence 161, App
19	18	10.2	323	21	US-10-617-978-14_COPY_64_240	Sequence 165, App
20	18	10.2	323	21	US-10-617-978-14_COPY_64_240	Sequence 169, App
21	18	10.2	3312	13	US-10-617-978-14_COPY_64_240	Sequence 113132, A
22	18	10.2	3312	13	US-10-617-978-14_COPY_64_240	Sequence 523, App
23	18	10.2	4790	17	US-10-617-978-14_COPY_64_240	Sequence 523, App
24	18	10.2	4790	17	US-10-617-978-14_COPY_64_240	Sequence 523, App
25	18	10.2	5587	17	US-10-617-978-14_COPY_64_240	Sequence 523, App
26	18	10.2	6846	20	US-10-617-978-14_COPY_64_240	Sequence 3132, App
27	18	10.2	6846	20	US-10-617-978-14_COPY_64_240	Sequence 2924, App
28	18	10.2	7037	15	US-10-617-978-14_COPY_64_240	Sequence 106, App
29	18	10.2	7037	15	US-10-617-978-14_COPY_64_240	Sequence 147, App
30	18	10.2	7037	15	US-10-617-978-14_COPY_64_240	Sequence 147, App
31	18	10.2	7132	13	US-10-617-978-14_COPY_64_240	Sequence 780, App
32	18	10.2	7974	19	US-10-617-978-14_COPY_64_240	Sequence 93, Appl
33	17	9.6	396	9	US-09-825-294-93	Sequence 5, Appl
34	17	9.6	396	9	US-09-825-294-93	Sequence 93, Appl
35	17	9.6	396	15	US-10-212-677-93	Sequence 93, Appl
36	17	9.6	396	17	US-10-361-811-93	Sequence 93, Appl
37	17	9.6	396	17	US-10-369-186-93	Sequence 93, Appl
38	17	9.6	427	19	US-10-674-124A-5885	Sequence 5885, App
39	17	9.6	468	15	US-10-338-694-25	Sequence 25, Appl
40	17	9.6	489	17	US-10-242-535A-27823	Sequence 27823, A
41	17	9.6	489	18	US-10-085-783A-27823	Sequence 27823, A
42	17	9.6	600	22	US-10-972-079-80648	Sequence 80648, A
43	17	9.6	758	9	US-09-910-943-462	Sequence 462, App
44	17	9.6	963	13	US-10-027-632-256780	Sequence 256780, A
45	17	9.6	963	17	US-10-027-632-256780	Sequence 256780, A

```
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(665)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-210235

Query Match      10.7%; Score 19; DB 13; Length 665;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 142 TGTGAATTTCTGAAGGATG 160
    |||||
Db 347 TGTGAATTTCTGAAGGATG 365

RESULT 2
US-10-027-632-210235
; Sequence 210235, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 210235
; LENGTH: 665
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(665)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-210235

Query Match      10.7%; Score 19; DB 17; Length 665;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 142 TGTGAATTTCTGAAGGATG 160
    |||||
Db 347 TGTGAATTTCTGAAGGATG 365

RESULT 3
US-10-087-192-997/c
; Sequence 997, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
```

```
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 997
; LENGTH: 143412
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(143412)
; OTHER INFORMATION: n = A,T,C or G
US-10-087-192-997

Query Match      10.7%; Score 19; DB 13; Length 143412;
Best Local Similarity 100.0%; Pred. No. 1.8;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 127 GCCTTCCAATGCTGGTGTG 145
    |||||
Db 129843 GCCTTCCAATGCTGGTGTG 129825

RESULT 4
US-10-721-793-139
; Sequence 139, Application US/10721793
; Publication No. US2005006531A1
; GENERAL INFORMATION:
; APPLICANT: Corona Villegas, Miguel
; APPLICANT: Garcia Rodriguez, Ma Consuelo
; APPLICANT: Valdez Cruz, Norma Adriana
; APPLICANT: Gurrola Briones, Georgina
; APPLICANT: Becerril Lujan, Baltazar
; APPLICANT: Posseani Postay, Lourival Domingos
; TITLE OF INVENTION: Recombinant immunogens for the Generation of Antivenoms to the
; FILE REFERENCE: 2099.0070001
; CURRENT APPLICATION NUMBER: US/10/721,793
; CURRENT FILING DATE: 2003-11-26
; PRIOR APPLICATION NUMBER: US 60/430,067
; PRIOR FILING DATE: 2002-12-02
; NUMBER OF SEQ ID NOS: 294
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 139
; LENGTH: 195
; TYPE: DNA
; ORGANISM: Centruroides sculpturatus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(195)
; OTHER INFORMATION: Product= Sodium-channel modifier toxin
; PUBLICATION INFORMATION:
; AUTHORS: Corona, M., Valdez-Cruz, N.A., Merino, E., Zurita, M. & Posseani L.D.
; TITLE: Genes and peptides from the scorpion Centruroides sculpturatus Ewing,
; JOURNAL: Toxicon
; VOLUME: 39
; ISSUE: 12
; PAGES: 1893-1898
; DATE: 2001-12-01
; DATABASE ENTRY DATE:
; RELEVANT RESIDUES: (1)..(195)
US-10-721-793-139

Query Match      10.2%; Score 18; DB 21; Length 195;
Best Local Similarity 100.0%; Pred. No. 6.1;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 111 TTACGGGTATTGTACGC 128
    |||||
```

Db 111 TTACGGGTATTGCTACGC 128

RESULT 5

US-10-721-793-143
; Sequence 143, Application US/10721793
; Publication No. US20050065331A1
; GENERAL INFORMATION:
; APPLICANT: Corona Villegas, Miguel
; APPLICANT: Garcia Rodriguez, Ma Consuelo
; APPLICANT: Valdez Cruz, Norma Adriana
; APPLICANT: Gurrola Briones, Georgina
; APPLICANT: Becerril Lujan, Baltazar
; APPLICANT: Possani Postay, Lourival Domingos
; TITLE OF INVENTION: Recombinant Immunogens for the Generation of Antivenoms to the
; TITLE OF INVENTION: Venom of Scorpions of the Genus Centruroides
; FILE REFERENCE: 2099.0070001
; CURRENT APPLICATION NUMBER: US/10/721,793
; CURRENT FILING DATE: 2003-11-26
; PRIOR APPLICATION NUMBER: US 60/430,067
; PRIOR FILING DATE: 2002-12-02
; NUMBER OF SEQ ID NOS: 294
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 143
; LENGTH: 195
; TYPE: DNA
; ORGANISM: Centruroides sculpuratus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(195)
; OTHER INFORMATION: Product= Sodium-channel modifier toxin
; PUBLICATION INFORMATION:
; AUTHORS: Corona, M., Valdez-Cruz, N.A., Merino, E., Zurita, M. & Possani L.D.
; TITLE: Genes and peptides from the scorpion Centruroides sculpuratus Ewing,
; JOURNAL: Toxicon
; VOLUME: 39
; ISSUE: 12
; PAGES: 1893-1898
; DATE: 2001-12-01
; DATABASE ENTRY DATE:
; RELEVANT RESIDUES: (1)..(195)
US-10-721-793-143

Query Match 10.2%; Score 18; DB 21; Length 195;

Best Local Similarity 100.0%; Pred. No. 6.1; Mismatches 0; Indels 0; Gaps 0;
Matches 18; Conservative 0;

Oy 111 TTACGGGTATTGCTACGC 128

Db 111 TTACGGGTATTGCTACGC 128

RESULT 6

US-10-721-793-147
; Sequence 147, Application US/10721793
; Publication No. US20050065331A1
; GENERAL INFORMATION:
; APPLICANT: Corona Villegas, Miguel
; APPLICANT: Garcia Rodriguez, Ma Consuelo
; APPLICANT: Valdez Cruz, Norma Adriana
; APPLICANT: Gurrola Briones, Georgina
; APPLICANT: Becerril Lujan, Baltazar
; APPLICANT: Possani Postay, Lourival Domingos
; TITLE OF INVENTION: Recombinant Immunogens for the Generation of Antivenoms to the
; TITLE OF INVENTION: Venom of Scorpions of the Genus Centruroides
; FILE REFERENCE: 2099.0070001
; CURRENT APPLICATION NUMBER: US/10/721,793
; CURRENT FILING DATE: 2003-11-26
; PRIOR APPLICATION NUMBER: US 60/430,067
; PRIOR FILING DATE: 2002-12-02
; NUMBER OF SEQ ID NOS: 294
; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 147
; LENGTH: 195
; TYPE: DNA
; ORGANISM: Centruroides sculpuratus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(195)
; OTHER INFORMATION: Product= Sodium-channel modifier toxin
; PUBLICATION INFORMATION:
; AUTHORS: Corona, M., Valdez-Cruz, N.A., Merino, E., Zurita, M. & Possani L.D.
; TITLE: Genes and peptides from the scorpion Centruroides sculpuratus Ewing,
; JOURNAL: Toxicon
; VOLUME: 39
; ISSUE: 12
; PAGES: 1893-1898
; DATE: 2001-12-01
; DATABASE ENTRY DATE:
; RELEVANT RESIDUES: (1)..(195)
US-10-721-793-147

Query Match 10.2%; Score 18; DB 21; Length 195;

Best Local Similarity 100.0%; Pred. No. 6.1; Mismatches 0; Indels 0; Gaps 0;
Matches 18; Conservative 0;

Oy 111 TTACGGGTATTGCTACGC 128

Db 111 TTACGGGTATTGCTACGC 128

RESULT 7

US-10-721-793-167
; Sequence 167, Application US/10721793
; Publication No. US20050065331A1
; GENERAL INFORMATION:
; APPLICANT: Corona Villegas, Miguel
; APPLICANT: Garcia Rodriguez, Ma Consuelo
; APPLICANT: Valdez Cruz, Norma Adriana
; APPLICANT: Gurrola Briones, Georgina
; APPLICANT: Becerril Lujan, Baltazar
; APPLICANT: Possani Postay, Lourival Domingos
; TITLE OF INVENTION: Recombinant Immunogens for the Generation of Antivenoms to the
; TITLE OF INVENTION: Venom of Scorpions of the Genus Centruroides
; FILE REFERENCE: 2099.0070001
; CURRENT APPLICATION NUMBER: US/10/721,793
; CURRENT FILING DATE: 2003-11-26
; PRIOR APPLICATION NUMBER: US 60/430,067
; PRIOR FILING DATE: 2002-12-02
; NUMBER OF SEQ ID NOS: 294
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 167
; LENGTH: 195
; TYPE: DNA
; ORGANISM: Centruroides sculpuratus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(195)
; OTHER INFORMATION: Product= Sodium-channel modifier toxin
; PUBLICATION INFORMATION:
; AUTHORS: Corona, M., Valdez-Cruz, N.A., Merino, E., Zurita, M. & Possani L.D.
; TITLE: Genes and peptides from the scorpion Centruroides sculpuratus Ewing,
; JOURNAL: Toxicon
; VOLUME: 39
; PAGES: 1893-1898
; DATE: 2001-12-01
; DATABASE ENTRY DATE:
; RELEVANT RESIDUES: (1)..(195)
US-10-721-793-167

Query Match

10.2%; Score 18; DB 21; Length 195;

Best Local Similarity 100.0%; Pred. No. 6.1;

```
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 111 TTACGGGTATTGCTACGC 128
Db 111 TTACGGGTATTGCTACGC 128

RESULT 8
US-10-721-793-171
; Sequence 171, Application US/10721793
; Publication No. US20050065331A1
; GENERAL INFORMATION:
; APPLICANT: Corona Villegas, Miguel
; APPLICANT: Garcia Rodriguez, Ma Consuelo
; APPLICANT: Valdez Cruz, Norma Adriana
; APPLICANT: Gurrola Briones, Georgina
; APPLICANT: Becerril Lujan, Baltazar
; APPLICANT: Possani Postay, Lourival Domingos
; TITLE OF INVENTION: Recombinant Immunogens for the Generation of Antivenoms to the
; FILE REFERENCE: 2099.0070001
; CURRENT APPLICATION NUMBER: US/10/721,793
; PRIOR FILING DATE: 2003-11-26
; PRIOR FILING DATE: 2002-12-02
; NUMBER OF SEQ ID NOS: 294
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 171
; LENGTH: 195
; TYPE: DNA
; ORGANISM: Centruroides sculpturatus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(195)
; OTHER INFORMATION: Product= Sodium-channel modifier toxin
; AUTHORS: Corona, M., Valdez-Cruz, N.A., Merino, E., Zurita, M. & Possani L.D.
; TITLE: Genes and peptides from the scorpion Centruroides sculpturatus Ewing,
; JOURNAL: Toxicon
; VOLUME: 39
; ISSUE: 12
; PAGES: 1893-1898
; DATE: 2001-12-01
; DATABASE ENTRY DATE:
; RELEVANT RESIDUES: (1)..(195)
;
Query Match 10.2%; Score 18; DB 21; Length 195;
Best Local Similarity 100.0%; Pred. No. 6.1;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 111 TTACGGGTATTGCTACGC 128
Db 111 TTACGGGTATTGCTACGC 128

RESULT 9
US-10-721-793-155
; Sequence 155, Application US/10721793
; Publication No. US20050065331A1
; GENERAL INFORMATION:
; APPLICANT: Corona Villegas, Miguel
; APPLICANT: Garcia Rodriguez, Ma Consuelo
; APPLICANT: Valdez Cruz, Norma Adriana
; APPLICANT: Gurrola Briones, Georgina
; APPLICANT: Becerril Lujan, Baltazar
; APPLICANT: Possani Postay, Lourival Domingos
; TITLE OF INVENTION: Recombinant Immunogens for the Generation of Antivenoms to the
; FILE REFERENCE: 2099.0070001
; CURRENT APPLICATION NUMBER: US/10/721,793
; PRIOR FILING DATE: 2003-11-26
; PRIOR FILING DATE: 2002-12-02
; NUMBER OF SEQ ID NOS: 294
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 155
; LENGTH: 198
; TYPE: DNA
; ORGANISM: Centruroides sculpturatus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(198)
; OTHER INFORMATION: Product= Sodium-channel modifier toxin
; AUTHORS: Corona, M., Valdez-Cruz, N.A., Merino, E., Zurita, M. & Possani L.D.
; TITLE: Genes and peptides from the scorpion Centruroides sculpturatus Ewing,
; JOURNAL: Toxicon
; VOLUME: 39
; ISSUE: 12
; PAGES: 1893-1898
; DATE: 2001-12-01
; DATABASE ENTRY DATE:
; RELEVANT RESIDUES: (1)..(198)
```

```
; PRIOR APPLICATION NUMBER: US 60/430,067
; PRIOR FILING DATE: 2002-12-02
; NUMBER OF SEQ ID NOS: 294
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 155
; LENGTH: 198
; TYPE: DNA
; ORGANISM: Centruroides sculpturatus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(198)
; OTHER INFORMATION: Product= Sodium-channel modifier toxin
; PUBLICATION INFORMATION:
; AUTHORS: Corona, M., Valdez-Cruz, N.A., Merino, E., Zurita, M. & Possani L.D.
; TITLE: Genes and peptides from the scorpion Centruroides sculpturatus Ewing,
; JOURNAL: Toxicon
; VOLUME: 39
; ISSUE: 12
; PAGES: 1893-1898
; DATE: 2001-12-01
; DATABASE ENTRY DATE:
; RELEVANT RESIDUES: (1)..(198)
;
US-10-721-793-155
Query Match 10.2%; Score 18; DB 21; Length 198;
Best Local Similarity 100.0%; Pred. No. 6.1;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 111 TTACGGGTATTGCTACGC 128
Db 111 TTACGGGTATTGCTACGC 128

RESULT 10
US-10-721-793-159
; Sequence 159, Application US/10721793
; Publication No. US20050065331A1
; GENERAL INFORMATION:
; APPLICANT: Corona Villegas, Miguel
; APPLICANT: Garcia Rodriguez, Ma Consuelo
; APPLICANT: Valdez Cruz, Norma Adriana
; APPLICANT: Gurrola Briones, Georgina
; APPLICANT: Becerril Lujan, Baltazar
; APPLICANT: Possani Postay, Lourival Domingos
; TITLE OF INVENTION: Recombinant Immunogens for the Generation of Antivenoms to the
; FILE REFERENCE: 2099.0070001
; CURRENT APPLICATION NUMBER: US/10/721,793
; CURRENT FILING DATE: 2003-11-26
; PRIOR APPLICATION NUMBER: US 60/430,067
; PRIOR FILING DATE: 2002-12-02
; NUMBER OF SEQ ID NOS: 294
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 159
; LENGTH: 198
; TYPE: DNA
; ORGANISM: Centruroides sculpturatus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(198)
; OTHER INFORMATION: Product= Sodium-channel modifier toxin
; PUBLICATION INFORMATION:
; AUTHORS: Corona, M., Valdez-Cruz, N.A., Merino, E., Zurita, M. & Possani L.D.
; TITLE: Genes and peptides from the scorpion Centruroides sculpturatus Ewing,
; JOURNAL: Toxicon
; VOLUME: 39
; ISSUE: 12
; PAGES: 1893-1898
; DATE: 2001-12-01
; DATABASE ENTRY DATE:
; RELEVANT RESIDUES: (1)..(198)
```

US-10-721-793-159

Query Match 10.2%; Score 18; DB 21; Length 198;
Best Local Similarity 100.0%; Pred. No. 6.1;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 111 TTACGGGTATTGCTACGC 128
||| ||||| ||||| |||||
Db 111 TTACGGGTATTGCTACGC 128

RESULT 11

US-10-721-793-163
; Sequence 163, Application US/10721793
; Publication No. US20050065331A1
; GENERAL INFORMATION:
; APPLICANT: Corona Villegas, Miguel
; APPLICANT: Garcia Rodriguez, Ma Consuelo
; APPLICANT: Valdez Cruz, Norma Adriana
; APPLICANT: Gurrola Briones, Georgina
; APPLICANT: Becerril Lujan, Baltazar
; APPLICANT: Possani Postay, Lourival Domingos
; TITLE OF INVENTION: Recombinant Immunogens for the Generation of Antivenoms to the
; TITLE OF INVENTION: Venom of Scorpions of the Genus Centruroides
; FILE REFERENCE: 2099.0070001
; CURRENT APPLICATION NUMBER: US/10/721,793
; CURRENT FILING DATE: 2003-11-26
; PRIOR APPLICATION NUMBER: US 60/430,067
; PRIOR FILING DATE: 2002-12-02
; NUMBER OF SEQ ID NOS: 294
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 163
; LENGTH: 198
; TYPE: DNA
; ORGANISM: Centruroides sculpturatus

FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(198)
OTHER INFORMATION: Product= Sodium-channel modifier toxin
PUBLICATION INFORMATION:
AUTHORS: Corona, M., Valdez-Cruz, N.A., Merino, E., Zurita, M. & Possani L.D.
TITLE: Genes and peptides from the scorpion Centruroides sculpturatus Ewing,
TITLE: that recognize Na+-channels
JOURNAL: Toxicon
VOLUME: 39
ISSUE: 12
PAGES: 1893-1898
DATE: 2001-12-01
DATABASE ENTRY DATE:
RELEVANT RESIDUES: (1)..(198)

US-10-721-793-163

Query Match 10.2%; Score 18; DB 21; Length 198;
Best Local Similarity 100.0%; Pred. No. 6.1;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 111 TTACGGGTATTGCTACGC 128
||| ||||| ||||| |||||
Db 111 TTACGGGTATTGCTACGC 128

RESULT 12

US-10-424-599-13962/c
; Sequence 13962; Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B

; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 13962
; LENGTH: 272
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_112616C.1
US-10-424-599-13962

Query Match 10.2%; Score 18; DB 18; Length 272;
Best Local Similarity 100.0%; Pred. No. 6.2;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 137 GCTGGTGTGAATTTCTGA 154
||| ||||| ||||| |||||
Db 96 GCTGGTGTGAATTTCTGA 79

RESULT 13

US-10-721-793-153
; Sequence 153, Application US/10721793
; Publication No. US20050065331A1
; GENERAL INFORMATION:
; APPLICANT: Corona Villegas, Miguel
; APPLICANT: Garcia Rodriguez, Ma Consuelo
; APPLICANT: Valdez Cruz, Norma Adriana
; APPLICANT: Gurrola Briones, Georgina
; APPLICANT: Becerril Lujan, Baltazar
; APPLICANT: Possani Postay, Lourival Domingos
; TITLE OF INVENTION: Recombinant Immunogens for the Generation of Antivenoms to the
; TITLE OF INVENTION: Venom of Scorpions of the Genus Centruroides
; FILE REFERENCE: 2099.0070001
; CURRENT APPLICATION NUMBER: US/10/721,793
; CURRENT FILING DATE: 2003-11-26
; PRIOR APPLICATION NUMBER: US 60/430,067
; PRIOR FILING DATE: 2002-12-02
; NUMBER OF SEQ ID NOS: 294
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 153
; LENGTH: 319
; TYPE: DNA
; ORGANISM: Centruroides sculpturatus

FEATURE:

NAME/KEY: CDS

LOCATION: (1)..(261)

OTHER INFORMATION: Product= Sodium-channel modifier toxin precursor

OTHER INFORMATION: In the mature peptide, the last 2 basic aminoacids are cut

FEATURE:

NAME/KEY: 3'UTR

LOCATION: (265)..(319)

OTHER INFORMATION:

FEATURE:

NAME/KEY: mat_peptide

LOCATION: (58)..(1)

OTHER INFORMATION:

FEATURE:

NAME/KEY: sig_peptide

LOCATION: (1)..(57)

OTHER INFORMATION:

PUBLICATION INFORMATION:

AUTHORS: Corona, M., Valdez-Cruz, N.A., Merino, E., Zurita, M. & Possani L.D.

TITLE: Genes and peptides from the scorpion Centruroides sculpturatus Ewing,

TITLE: that recognize Na+-channels

JOURNAL: Toxicon

VOLUME: 39

ISSUE: 12

PAGES: 1893-1898

DATE: 2001-12-01

DATABASE ENTRY DATE:

RELEVANT RESIDUES: (1)..(261)

US-10-721-793-153

```
Query Match      10.2%; Score 18; DB 21; Length 319;
Best Local Similarity 100.0%; Pred. No. 6.2;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Qy 111 TTACGGGTATTGCTACGC 128
Db 168 TTACGGGTATTGCTACGC 185

```

RESULT 14
US-10-721-793-137
; Sequence 137, Application US/10721793
; Publication No. US20050065331A1
; GENERAL INFORMATION:
; APPLICANT: Corona Villegas, Miguel
; APPLICANT: Garcia Rodriguez, Ma Consuelo
; APPLICANT: Valdez Cruz, Norma Adriana
; APPLICANT: Gurrola Briones, Georgina
; APPLICANT: Becerril Lujan, Baltazar
; APPLICANT: Possani Postay, Lourival Domingos
; TITLE OF INVENTION: Recombinant Immunogens for the Generation of Antivenoms to the
; TITLE OF INVENTION: Venom of Scorpions of the Genus Centruroides
; FILE REFERENCE: 2099.0070001
; CURRENT APPLICATION NUMBER: US/10721.793
; CURRENT FILING DATE: 2003-11-26
; PRIOR APPLICATION NUMBER: US 60/430,067
; PRIOR FILING DATE: 2002-12-02
; NUMBER OF SEQ ID NOS: 294
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 137
; LENGTH: 323
; TYPE: DNA
; ORGANISM: Centruroides sculpturatus
; FEATURE:
; NAME/KEY: CD3
; LOCATION: (5)..(265)
; OTHER INFORMATION: Product= Sodium-channel modifier toxin precursor
; OTHER INFORMATION: In the mature peptide, the last Cys is amidated, and the last Gly
; OTHER INFORMATION: and the last 2 basic aminoacids are cut
; FEATURE:
; NAME/KEY: 3'UTR
; LOCATION: (249)..(323)
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: 5'UTR
; LOCATION: (1)..(4)
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: (64)..(1)
; OTHER INFORMATION: Product= Sodium-channel modifier toxin
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: (5)..(61)
; OTHER INFORMATION:
; PUBLICATION INFORMATION:
; AUTHORS: Corona, M., Valdez-Cruz, N.A., Merino, E., Zurita, M. & Possani L.D.
; TITLE: Genes and peptides from the scorpion Centruroides sculpturatus Ewing,
; TITLE: that recognize Na+-channels
; JOURNAL: Toxicon
; VOLUME: 39
; ISSUE: 12
; PAGES: 1893-1898
; DATE: 2001-12-01
; DATABASE ENTRY DATE:
; RELEVANT RESIDUES: (5)..(265)
US-10-721-793-137

```

QY 111 TTACGGGTATTGCTACGC 128
|||
Db 172 TTACGGGTATTGCTACGC 189

```

RESULT 15
US-10-721-793-141
; Sequence 141, Application US/10721793
; Publication No. US20050065331A1
; GENERAL INFORMATION:
; APPLICANT: Corona Villegas, Miguel
; APPLICANT: Garcia Rodriguez, Ma Consuelo
; APPLICANT: Valdez Cruz, Norma Adriana
; APPLICANT: Gueroa Briones, Georgina
; APPLICANT: Becerril Lujan, Baltazar
; APPLICANT: Possani Postay, Lourival Domingos
; TITLE OF INVENTION: Recombinant Immunogens for the Generation of Antivenoms to the
; TITLE OF INVENTION: Venom of Scorpions of the Genus Centruroides
; FILE REFERENCE: 2099.0070001
; CURRENT APPLICATION NUMBER: US/10/721,793
; CURRENT FILING DATE: 2003-11-26
; PRIOR APPLICATION NUMBER: US 60/430,067
; PRIOR FILING DATE: 2002-12-02
; NUMBER OF SEQ ID NOS: 294
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 141
; LENGTH: 323
; TYPE: DNA
; ORGANISM: Centruroides sculptratus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (5)..(265)
; OTHER INFORMATION: Product= Sodium-channel modifier toxin precursor
; OTHER INFORMATION: In the mature peptide, the last Cys is amidated, and the last Gly
; OTHER INFORMATION: and the last 2 basic aminoacids are cut
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: (62)..( )
; OTHER INFORMATION: Product= Sodium-channel modifier toxin
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: (5)..(61)
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: 3'UTR
; LOCATION: (269)..(323)
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: 5'UTR
; LOCATION: (1)..(4)
; OTHER INFORMATION:
; PUBLICATION INFORMATION:
; AUTHORS: Corona, M., Valdez-Cruz, N.A., Merino, E., Zurita, M. & Possani L.D.
; TITLE: Genes and peptides from the scorpion Centruroides sculptratus Ewing,
; TITLE: that recognize Nav-channels
; JOURNAL: Toxicon
; VOLUME: 39
; ISSUE: 12
; PAGES: 1893-1898
; DATE: 2001-12-01
; DATABASE ENTRY DATE:
; RELEVANT RESIDUES: (5)..(265)
US-10-721-793-141

```

QY 111 TTACGGGTATTGCTACGC 128
|||
Db 172 TTACGGGTATTGCTACGC 189

Search completed: August 26, 2005, 21:08:04
Job time : 415.026 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 26, 2005, 20:14:05 ; Search time 2114.25 Seconds
(without alignments)
3186.647 Million cell updates/sec

Title: US-10-617-978-14_COPY_64_240

Perfect score: 177

Sequence: 1 gctgacgtccgggaacta.....atgagaacgtgaaggtctga 177

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 34239544 seqs, 19032134700 residues

Word size : 0

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : EST.*

- 1: gb_est1.*
2: gb_est2.*
3: gb_hcc.*
4: gb_est3.*
5: gb_est4.*
6: gb_est5.*
7: gb_est6.*
8: gb_gsa1.*
9: gb_gsa2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
C 1	22	12.4	435	2	BE950396	UI-M-CE0
C 2	21	11.9	379	7	CK817241	UMC-bov 0
C 3	21	11.9	665	1	AJ814218	AJ814218
C 4	21	11.9	674	6	CB166163	KXB603014
C 5	21	11.9	728	7	CK949817	4074871 B
C 6	19	10.7	580	2	AW361823	PM0-CT026
C 7	19	10.7	1101	8	B10170	F14D20-SP6
C 8	18	10.2	263	2	BS586775	BS586775
C 9	18	10.2	313	4	BG603087	EST502177
C 10	18	10.2	360	8	BZ686719	PUBBU54TD
C 11	18	10.2	423	8	AQ405655	HS 5049 B
C 12	18	10.2	466	7	CF794539	890065 MA
C 13	18	10.2	466	8	A2644277	1M0508H11
C 14	18	10.2	477	4	BG603242	EST502332
C 15	18	10.2	483	9	PT003115U	Parameciu
C 16	18	10.2	547	8	BZ675947	PUBGF58TD
C 17	18	10.2	554	5	BP615364	BP615364
C 18	18	10.2	587	9	CG900317	ZMMBB023
C 19	18	10.2	589	8	AQ350422	RPCI11-1
C 20	18	10.2	610	4	BG603086	EST502176
C 21	18	10.2	611	7	CK898930	SGF162516
C 22	18	10.2	621	7	CF793169	886193 MA
C 23	18	10.2	623	6	CB963735	AGENCOURT
C 24	18	10.2	628	4	BJ623886	BJ623886

25	18	10.2	640	8	AZ710307	RPCI-24-1
26	18	10.2	656	2	BF611918	de89h10.Y
27	18	10.2	675	5	BX850657	BX850657
28	18	10.2	724	6	CB592178	AGENCOURT
29	18	10.2	732	9	BX130112	Danio rer
30	18	10.2	748	7	CK352152	AGENCOURT
31	18	10.2	752	9	CC923223	t069n21ba
32	18	10.2	755	5	BP678876	BP678876
33	18	10.2	756	6	CB206666	AGENCOURT
34	18	10.2	781	4	BG400860	602464085
35	18	10.2	795	5	B0910572	AGENCOURT
36	18	10.2	806	8	BH360033	CH230-18J
37	18	10.2	806	8	BH360034	CH230-18J
38	18	10.2	819	9	CG791244	ZMBBB028
39	18	10.2	821	9	CC490159	CH240 323
40	18	10.2	822	9	CG043458	PUIGE43TD
41	18	10.2	857	5	BQ736159	AGENCOURT
42	18	10.2	869	6	CB941331	AGENCOURT
43	18	10.2	948	6	CD255176	AGENCOURT
44	18	10.2	1055	4	BM802977	AGENCOURT
45	18	10.2	1084	6	CA974306	AGENCOURT

ALIGNMENTS

RESULT 1
BE950396/c
LOCUS
DEFINITION
UI-M-CE0-ayw-a-10-0-UI.s1 NIH BMAP Ret3 Mus musculus CDNA clone
UI-M-CE0-ayw-a-10-0-UI 3', mRNA sequence.
ACCESSION
BE950396
VERSION
BE950396.1 GI:10589062
KEYWORDS
EST.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 435)
Bonaldo, M.F., Lennon, G. and Soares, M.B.
REFERENCE
Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)
JOURNAL
97044477
MEDLINE
8889548
PUBMED
COMMENT
Contact: Chin, H
National Institute of Mental Health
6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD
20892-9643, USA
Tel: 301 443 1706
Fax: 301 443 9890
Email: mEST@mail.nih.gov
The sequence contained an oligo-dT track that was present in the
oligonucleotide that was used to prime the synthesis of first
strand cDNA and therefore this may represent a bonafide poly A
tail. The sequence tag present in the cDNA between the NotI site
and the oligo-dT track served to verify it as a clone from the
retina tissue cDNA Library Preparation: M.B. Soares Lab Clone
distribution: Researchers may obtain BMAP CDNA clones from RESEARCH
GENETICS. It should be noted that Bento Soares is generating a
small number of additional specialized non-redundant arrays of BMAP
cDNAs whose availability will be considered under appropriate and
limited collaborative arrangements. The tissue for this library was
contributed by Dr. Xin-Yuan Fu, Yale University School of Medicine
The following repetitive elements were found in this cDNA sequence:
1-21, >AT richlow complexity 169-428, >Lx9#LINE/L1
Seq primer: M13 Forward
POLYA=Yes.

FEATURES
source
Location/Qualifiers
1. .435
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"

```

/db_xref="taxon:10090"
/clone="UI-M-CE0-ayw-a-10-0-UI"
/dev_stages="6 weeks"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NIH BMAP Ret3"
/notes="Vector: pVT3D-Pac (Pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: Eco RI; The
NIH_BMAP_Ret3 library is derived from mouse retina tissue.
For a detailed description of the library from which this
clone was derived, please visit our web site at
braineest.eng.uiowa.edu. The tissue for this library was
contributed by Dr. Xin-Yuan Fu, Yale University School of
Medicine
TAG_TISSUE=adult-retina
TAG_LIB=NIH BMAP Ret3
TAG_SEQ=GTGAGCGGCAC"

ORIGIN

Query Match      12.4%; Score 22; DB 2; Length 435;
Best Local Similarity 100.0%; Pred. No. 0.41;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 81 CATTAAAGATCTGTGAGAAACAC 102
   |||||
Db 127 CATTAAAGATCTGTGAGAAACAC 106

RESULT 2
CK817241
LOCUS      379 bp      mRNA      linear      EST 01-MAR-2004
DEFINITION UMC-boy_0801-009-fl1 Day 0 Oviduct post-LH surge bov Bos taurus
           CDNA 3'; mRNA sequence.
ACCESSION  CK817241
VERSION    CK817241.1 GI:44834166
KEYWORDS   EST.
SOURCE     Bos taurus (cow)
ORGANISM   Bos taurus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
            Bovinae; Bos.
REFERENCE  1 (bases 1 to 379)
            Prather,R.S., Antoniou,E., Garverick,H.A., Green,J.A., Lucy,M.C.,
            Roberts,R.M., Smith,M.F. and Youngquist,R.S.
            USDA Grant NRI-2002-03476: Bovine ESTs: Focus on Female
            Reproduction
JOURNAL    Unpublished (2002)
COMMENT    Contact: DNA Core Facility (Bovine Project)
            Animal Science - RS Prather
            University of Missouri-Columbia
            M616 Medical Sciences Bldg., Columbia, MO 65212, USA
            Tel: (573)882-0428
            Fax: (573)884-5552
            Email: bovine@rnet.missouri.edu
            POLYA=Yes.

FEATURES             Location/Qualifiers
     1..379
         /organism="Bos taurus"
         /mol_type="mRNA"
         /db_xref="taxon:9913"
         /clone_lib="bov"
     /note="Funding: The production of ESTs submitted in this
     project was funded by USDA Grant MRI-2002-03476 entitled
     'Bovine ESTs: Focus on Female Reproduction' to RS Prather,
     E Antoniou, HA Garverick, JA Green, MC Lucy, RM Roberts,
     MF Smith and RS Youngquist. Genetic Source: Heifers for
     the project were purchased from Circle A Ranch, Iberia, MO
     (http://www.circlearanch.com/home.html). These heifers,
     while not registered have known Angus pedigrees going back
     at least 4 generations. Samples collected: The samples
     consisted of the following: germinal vesicle-stage
     oocytes; in vitro derived embryos (2-cell, morula,
     blastocyst and nuclear transfer blastocyst); in vivo
     blastocysts and conceptuses (days 8, 14, 16 and 18);

```

corpora lutea (days 3, 5, 8, 14, 16, 18 and 35); ovarian follicles (days 0, non-recruited, recruited, early selected and preovulatory); oviduct (days 0, 3 and 5); endometrium (days 5, 8, 14, 16, 18 and 35); and placenta/embryo from day 35 conceptuses. Expanded descriptions of how the tissues were collected can be found at the following URL: <http://genome.rnet.missouri.edu/Bovine/Methods.html>. Library construction (Standard Protocol): All procedures have been described in detail elsewhere (Soares et al., 1994; Bonaldo et al., 1996; Jiang et al., 2001). Total cellular RNA from each sample was isolated by using STAT-60 reagent (Tel-Test, Friendswood, TX) and the poly(A)+ RNA was obtained by two rounds of purification with the Oligotex mRNA isolation kit (Qiagen) according to the manufacturer's instructions. The libraries were constructed by E. Ferguson and R. Woods essentially as described by the manufacturer's instructions provided with the SuperScript Plasmid System (Invitrogen, cat. no. 18248-013). Briefly, 1mcg of poly(A)+ RNA was annealed at 37 degrees with 10mcg of NotI-tag-dT18 oligonucleotide (GTGTCGCGCGCGC-tag-T18) and reverse transcribed at 37 degrees with SuperScript II (Invitrogen) reverse transcriptase (Jiang et al., 2001). The 'tag' represents a tissue/stage-specific ten-base sequence identifier (<http://genome.uiowa.edu/pubsoft/software.html>) present in the oligonucleotide used to prime first-strand synthesis. Second strand synthesis was performed with T4 DNA polymerase in the presence of DNA ligase and RNase H. After second strand synthesis, the double-stranded cDNAs were ligated to SalI adapters (Invitrogen-Life Technologies) and digested with NotI. The cDNAs were size selected by passage through cDNA size fractionation columns (Invitrogen-Life Technologies). The cDNAs derived from each developmental stage of a particular tissue were mixed on an equimolar basis and ligated directionally into the NotI and SalI sites of the pCMV-SPORT6 vector (Invitrogen). After ligation of the inserts, the plasmids were electroporated into DH10B bacteria. Library Construction (PCR Protocol): The amount of mRNA that was recovered from oocytes and embryos was quite limiting and was not sufficient for library production with the standard protocol. Therefore, PCR-based protocol was utilized for producing libraries from sources in which the amount of extracted mRNA was small (oocytes and embryos). Poly-A RNA was isolated by using the MicroPoly(A) Pure kit from Ambion (cat. # 1918). The mRNA was reverse transcribed with a NotI-tag-dT18 oligonucleotide and a SMART oligonucleotide (Clontech) modified to contain a SalI site to generate full-length cDNA with a sequence complementary to the SMART oligonucleotide. Sequences within the SMART and dT oligonucleotides were used as primers to amplify the cDNAs by PCR with pfu turbo polymerase (Stratagene). The resulting PCR products were purified, digested with NotI and SalI and size fractionated by using Chroma Spin-1000 columns (Clontech). Purified cDNA from each PCR reaction was quantitated and mixed on an equimolar basis for ligation into the pCMV-SPORT6 vector. Preliminary Library Characterization: Randomly chosen clones from each library were analyzed by restriction digestion to determine average insert size (96 clones) and by sequencing (~4 96-well plates) to confirm library quality [e.g. the presence of short poly(A) tails, genomic DNA contamination (must be <1%), ribosomal RNA clones (must be <1%), etc.] and to provide a sequence database representing the predominant clones in each library. The clones were sequenced at the University of Missouri-Columbia DNA Core Facility. After production of the libraries, equal numbers of recombinants from each library were pooled to produce a single mixed library (mega-library) for more extensive sequencing. Bioinformatics work was performed by GK Springer's bioinformatics group (Spollen WG, Topinka CM, Khambati AA)

in Computer Science at the University of Missouri-Columbia. Clone Requests: Requests for clones should be made to the Director of the University of Missouri DNA Core facility at: bovine@net.missouri.edu. Ronaldo MF, Lennon G, Soares MB. Normalization and Subtraction: Two approaches to facilitate gene discovery. Genome Res, 1996; 6:791-806. Jiang H, Bivens NJ, Ries JE, Whitworth KM, Green JA, Forrester LJ, Springer GK, Didion BA, Mathialagan N, Prather RS, Lucy MC (2001) Constructing cDNA libraries with fewer clones that contain long poly(dA) tails. Biotechniques 31:38-42. Soares MB, MF Banaldo, P Jelene, L Su, L Lawton, A Efstratiadis. 1994. Construction and characterization of a normalized cDNA library. Proc Natl Acad Sci, 91:9228-9232. TAG TISSUE=Day 0 Ovicult post-LH surge TAG_SEQ=TGTACCGATG

ORIGIN

Query Match : 11.9%; Score 21; DB 7; Length 379;
Best Local Similarity 100.0%; Pred. No. 1.5;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 81 CATTAAGATCTGTCAGAAACA 101
|||||
Db 343 CATTAAGATCTGTCAGAAACA 363

RESULT 3
AJ814218/c AJ814218 665 bp mRNA linear EST 13-SEP-2004
LOCUS AJ814218 KN206 Bos sp. cDNA clone C000519812, mRNA sequence.
DEFINITION
ACCESSION AJ814218
VERSION AJ814218.1 GI:51881694
KEYWORDS EST.
SOURCE Bos sp.
ORGANISM Bos sp.

REFERENCE 1 (bases 1 to 665)
AUTHORS McGuire,K. and Glass,E.J.
TITLE Unpublished ESTs, McGuire and Glass
JOURNAL Unpublished (2004)
COMMENT Contact: McGuire K
Genomics and Genetics
Roslin Institute
Roslin, Midlothian, EH25 9PS, UNITED KINGDOM
Single pass sequencing. Bases called and trimmed with phred v0.020425.c. Vector identified by cross match with the -minscore 20 and -minmatch 12 options. Vector:pBluescriptII(SK+) R. Site 1:EcoRV(lost) R. Site 2:NotI Seq Primer: T7 Normalised library constructed from pooled monocytes from Bos taurus (Holstein) and Bos indicus (Sahiwal) cattle subjected to various stimuli, including infection with the protozoan parasite Theileria annulata.

FEATURES

source
1 . .665
Location/Qualifiers
/organism="Bos sp."
/mol_type="mRNA"
/db_xref="taxon:29061"
/clone="C000519812"
/tissue_type="blood"
/cell_type="bovine monocyte"
/clone_lib="KN206"
/note="Vector: pBluescriptII(SK+); Site_1: EcoRV(lost); Site_2: NotI; Normalised library constructed from pooled monocytes from Bos taurus (Holstein) and Bos indicus (Sahiwal) cattle subjected to various stimuli, including infection with the protozoan parasite Theileria annulata"

ORIGIN

Query Match 11.9%; Score 21; DB 1; Length 665;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy

81 CATTAAGATCTGTCAGAAACA 101
|||||
Db 360 CATTAAGATCTGTCAGAAACA 340

RESULT 4

CB166163
LOCUS CB166163
DEFINITION KX6503014860.R1 CSQFXN32 hypothalamus Bos taurus cdna, mRNA sequence.
ACCESSION CB166163
VERSION CB166163.1 GI:28152289
KEYWORDS EST.
SOURCE Bos taurus (cow)
ORGANISM Bos taurus

ORIGIN

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovinae; Bos.
REFERENCE 1 (bases 1 to 674)
AUTHORS Adelson,D.L. and Gill,C.A.
TITLE Bovine ESTs (Adelson and Gill)
JOURNAL Unpublished (2003)
COMMENT Contact: David L. Adelson
Animal Breeding and Genetics
Texas A&M University
Animal Science Dept., TAMU-2471, College Station, TX 77843-2471, USA

Tel: 9798452616
Fax: 9798456970
Email: david.adelson@tamu.edu.
Location/Qualifiers
1 . .674
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
/tissue_type="hypothalamus"
/clone_lib="CSEQFXN32 hypothalamus"
/note="Torgan: hypothalamus; Vector: pBluescript SK+;
Site_1: NotI; Site_2: EcoRI; sequence 5' of the insert (5'-NNN...NNNinsert)
GCCGAATTGGAGCTCCACCGGTGGCGCGCGCTCGAG. Sequence 3' of the inserts (AAGAATTCGATATCAAGCTATTCGATACCGTCGACCTCGAG. normalized rd 1 library, sequenced 3' with M13R primer."

FEATURES

source
11.9%; Score 21; DB 6; Length 674;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

ORIGIN

Query Match 11.9%; Score 21; DB 6; Length 674;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 81 CATTAAGATCTGTCAGAAACA 101
|||||
Db 345 CATTAAGATCTGTCAGAAACA 365

RESULT 5

CK949817/c CK949817 728 bp mRNA linear EST 15-MAR-2004
LOCUS CK949817 BARC 10BOV Bos taurus cDNA clone 10BOV26_I08 5', mRNA sequence.
DEFINITION
ACCESSION CK949817
VERSION CK949817.1 GI:45464197
KEYWORDS EST.
SOURCE Bos taurus (cow)
ORGANISM Bos taurus

ORIGIN

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovinae; Bos.
REFERENCE 1 (bases 1 to 728)
AUTHORS Sonstegard, T.S., Van Tassel, C.P., Matukumalli, L.K., Harhay, G.P., Bosak, S., Rubenfield, M. and Gasbarre, L.C.
TITLE Production of EST from cDNA libraries derived from immunologically activated bovine gut

```

JOURNAL
COMMENT
Unpublished (2004)
Contáct: Tad S. Sonstegard
Bovine Functional Genomics Laboratory
Animal and Natural Resources Institute
Bldg# 200 Rm2A BARC-East, Beltsville, MD 20705, USA
Tel: 3015048416
Fax: 3015048414
Email: tads@anri.barc.usda.gov
Single pass sequencing. Bases called and trimmed with phred
0.000925 using options -trim_alt "-trim_fasta. Vector identified
by cross_match using options -minmatch 12 -minscore 12
Plate: 26 row: I column: 08
Seq primer: CCCAGTCACGACGTGTGAACAG
High quality sequence stop: 728.

FEATURES
source
Location/Qualifiers
1..728
/organism="Bos taurus"
/mol_type="mRNA"
/strain="Holstein"
/db_xref="taxon:9913"
/clone="10BOV26_108"
/sex="Male"
/tissue_type="Pooled"
/dev_stage="Multiple"
/lab_host="DH10B T1 phage resistant"
/clone_lib="BARC 108GV"
/notes="Organ: Small Intestine; Vector: pagen-1; Site_1:
EcoRI; Site_2: NotI; Equimolar amounts of mRNA extracted
from proximal jejunums of 18 and 21 wk old steers, and
distal ileums of 14 day old calves. proximal jejunum
exposed to C. oncophora for 3 and 6 weeks, and distal
ileum exposed to C. parvum for 7 days"

ORIGIN
Query Match 11.9%; Score 21; DB 7; Length 728;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 81 CATTAAGATCTGTGAGAAACA 101
|||||
Db 499 CATTAAGATCTGTGAGAAACA 479

RESULT 6
AW361823
LOCUS
DEFINITION
PMO-CT0263-021299-006-b02 CT0263 Homo sapiens cDNA, mRNA sequence.
ACCESSION
AW361823
VERSION
AW361823.1 GI:6866473
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (Bases 1 to 580)
HCCP http://www.ludwig.org.br/ORESTES.
The FAPESP/LiCR Human Cancer Genome Project
Unpublished (1999)
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LiCR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?ti=PM0&t2=PM0-CT0263-021299-006-b02&t3=1999-12-02&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 3
High quality sequence stop: 148.

FEATURES
source
Location/Qualifiers
1..580
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="CT0263"
/notes="Organ: colon; Vector: puc18; Site_1: SmaI; Site_2:
SmaI; A mini-library was made by cloning products derived
from ORESTES PCR (U.S. Letters Patent application No.
196,716 - Ludwig Institute for Cancer Research) profiles
into the pUC 18 vector. Reverse transcription of tissue
mRNA and cDNA amplification were performed under low
stringency conditions."

ORIGIN
Query Match 10.7%; Score 19; DB 2; Length 580;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 129 CTTCCAATGCTGTGTGAA 147
|||||
Db 98 CTTCCAATGCTGTGTGAA 116

RESULT 7
B10170
LOCUS
DEFINITION
F14D20-Sp6 IGF Arabidopsis thaliana genomic clone F14D20, genomic
survey sequence.
ACCESSION
B10170
VERSION
B10170.1 GI:2091289
KEYWORDS
GSS.
ORGANISM
Arabidopsis thaliana (thale cress)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
1 (Bases 1 to 1101)
PENG,J., Dewar,K., Buehler,E., Kim,C., Li,Y., Shinn,P., Sun,H. and
Ecker,J.
BAC End Sequences at ATGC
Unpublished (1997)
Other_GSSs: F14D20-T7
Contact: Ecker J.
Arabidopsis Thaliana Genome Center
University of Pennsylvania
Dept. of Biology, University of Pennsylvania, Philadelphia, PA
19104
Tel: 215-898-9384
Fax: 215-898-8780
Email: jecker@atgenome.bio.upenn.edu
Seq primer: Sp6
Class: BAC ends
High quality sequence start: 121
High quality sequence stop: 802.

FEATURES
source
Location/Qualifiers
1..1101
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/scotype="Columbia"
/db_xref="taxon:3702"
/clone="F14D20"
/sex="hermaphrodite"
/clone_lib="IGF"
/notes="Vector: BeloSACII; Site_1: EcoRI; Site_2: EcoRI;
Produced by Thomas Altmann"

ORIGIN
Query Match 10.7%; Score 19; DB 8; Length 1101;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 146 AATTTCTGAAGATGAGAA 164
 |||||
 Db 405 AATTTCTGAAGATGAGAA 423

RESULT 8
 BB586775 263 bp mRNA linear EST 30-NOV-2000
 LOCUS BB586775 RIKEN full-length enriched, adult male urinary bladder Mus
 DEFINITION musculus cDNA clone 9570005A14 5', mRNA sequence.

ACCESSION BB586775
 VERSION BB586775.1 GI:11483319
 KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM

Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

AUTHORS

Aizawa,K., Akahira,S., Akimura,T., Arai,A., Arakawa,T.,
 Carninci,P., Hanagaki,T., Hayatsu,N., Hiraoka,T., Hirozane,T.,
 Hodojima,Y., Imotani,K., Ishii,Y., Itoh,M., Izawa,M., Kawai,J.,
 Kojima,Y., Konno,H., Kuwabara,M., Matsuyama,T., Miyazaki,A.,
 Nakamura,M., Nishi,K., Nomura,K., Numazaki,K., Okazaki,Y.,
 Okubo,T., Owa,C., Sakai,C., Sakai,K., Sasaki,D., Sato,K.,
 Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T., Sogabe,Y.,
 Suzuki,H., Tagawa,A., Takahashi,F., Tanaka,T., Toya,T.,
 Watahiki,A., Yamamura,T., Yasunishi,A., Yoshida,K., Yoshiki,A.,
 Muramatsu,M. and Hayashizaki,Y.

RIKEN Mouse ESTs (Aizawa,K. et al. 2000)

TITLE

JOURNAL

COMMENT

Contact: Yoshihide Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic
 Sciences Center (GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-2 Suenhiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 Tel: 81-45-503-9222
 Fax: 81-45-503-9216

Email: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/
 Carninci,P., Nishiyama,Y., Westover,A., Itoh,M., Nagaoaka,S.,
 Sasaki,N., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
 Thermotabilization and thermoactivation of thermostable enzymes by
 trehalose and its application for the synthesis of full length
 cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)
 Itoh,M., Kitsuai,T., Akiyama,J., Shibata,K., Izawa,M., Kawai,J.,
 Tomaru,Y., Carninci,P., Shibata,Y., Ozawa,Y., Muramatsu,M.,
 Okazaki,Y. and Hayashizaki,Y.
 Automated filtration-based high-throughput plasmid preparation
 system. Genome Res. 9 (5), 463-470 (1999)
 Carninci,P. and Hayashizaki,Y.
 High-efficiency full-length cDNA cloning. Methods Enzymol. 303,
 19-44 (1999)

Please visit our web site (http://genome.rtc.riken.go.jp) for
 further details.

FEATURES

source

Location/Qualifiers

1..263
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /clone="9570005A14"
 /sex="male"
 /tissue_type="urinary bladder"
 /dev_stage="adult"
 /lab_host="DH10B"
 /clone_lib="RIKEN full-length enriched, adult male urinary
 bladder"

/note="Site 1: Sali; Site 2: BamHI; cDNA library was
 prepared and sequenced in Mouse Genome Encyclopedia
 Project of Genome Exploration Research Group in Riken
 Genomic Sciences Center and Genome Science Laboratory in
 RIKEN. Division of Experimental Animal Research in Riken
 contributed to prepare mouse tissues. 1st strand cDNA was
 primed with a primer [5'
 GAGAGAGAGAGGATCCAGAGCTCTTTTITTTTTTITVN 3'], cDNA was

prepared by using trehalose thermo-activated reverse
 transcriptase and subsequently enriched for full-length by
 cap-trapper. cDNA went through one round of normalization
 to Rot = 20.0 and subtraction to Rot = 370.0. Second
 strand cDNA was prepared with the primer adapter of
 sequence [5' GAGAGAGATTCGAGTTAATTAATATCCCCCCCCCCC
 3']. cDNA was cleaved with XhoI and BamHI. Vector: a
 modified pBluescript KS(+) after bulk excision from
 Lambda FLC I."

ORIGIN

Query Match 10.2%; Score 18; DB 2; Length 263;
 Best Local Similarity 100.0%; Pred. No. 76;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 142 TGTGAATTTCTGAAGGAT 159
 |||||

Db 168 TGTGAATTTCTGAAGGAT 185
 |||||

RESULT 9

LOCUS

DEFINITION

BB603087 313 bp mRNA linear EST 14-AUG-2001
 EST502177 Plasmodium yoelii sporozoite cDNA Plasmodium yoelii cDNA
 clone PYCDN74, mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

COMMENT

Plasmodium yoelii
 Plasmodium yoelii
 1 (bases 1 to 313)
 Kappe,S.H.I., Gardner,M.J., Brown,S.M., Ross,J., Matuschewski,K.,
 Ribeiro,J.M., Adams,J.H., Quackenbush,J., Cho,J., Carucci,D.J.,
 Hoffman,S.L. and Nussenzweig,V.
 Exploring the transcriptome of the malaria sporozoite stage
 Proc. Natl. Acad. Sci. U.S.A. 98 (17), 9895-9900 (2001)

21396555

11493895

Contact: Malcolm J. Gardner

Department of Eukaryotic Genomics

The Institute for Genomic Research

9712 Medical Center Drive, Rockville, MD 20850, USA

Tel: 301 838 3519

Fax: 301 838 0208

Email: gardner@tigr.org

Request for clones, please contact: Stefan Kappe,

kappes01@popmail.med.nyu.edu Michael Heidelberg Division,

Department of Pathology New York University School of Medicine.

Location/Qualifiers

1..313

/organism="Plasmodium yoelii"

/mol_type="mRNA"

/strain="17XNL"

/db_xref="taxon:5861"

/clone="PYCDN74"

/dev_stage="sporozoites from salivary gland"

/lab_host="E. coli TOP10"

/clone_lib="Plasmodium yoelii sporozoite cDNA"

/note="Vector: pCR4; TA cloning; Plasmodium yoelii
 sporozoite cDNA library from salivary gland sporozoites 14
 days post-infection"

ORIGIN

Query Match 10.2%; Score 18; DB 4; Length 313;
 Best Local Similarity 100.0%; Pred. No. 77;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 134 AATCGTGGTGAATTC 151
 |||||

Db 218 AATCGTGGTGAATTC 201
 |||||

```

RESULT 10
LOCUS      BZ686719/c
DEFINITION PUBBU54TD ZM_0.6_1.0_KB Zea mays genomic clone ZMMBTA019J11,
            genomic survey sequence.
ACCESSION  BZ686719
VERSION     BZ686719.1 GI:28247227
KEYWORDS    GSS.
SOURCE      Zea mays
ORGANISM    Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1 (bases 1 to 360)
AUTHORS     Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
            Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
            Bennetzen,J.
            Maize Genomics Consortium
            Unpublished (2003)
            Contact: Cathy Whitelaw
TITLE       TIGR
JOURNAL     9712 Medical Center Drive, Rockville, MD 20850, USA
COMMENT     Tel: 301-838-5843
            Fax: 301-838-0208
            Email: whitelaw@tigr.org
            Seq primer: TF
            Class: sheared ends.
FEATURES    source
            Location/Qualifiers
                1..360
                /organism="Zea mays"
                /mol_type="genomic DNA"
                /strain="B73"
                /db_xref="taxon:4577"
                /clone="ZMMBTA019J11"
                /clone_lib="ZM 0.6_1.0_KB"
                /notes="Vector: PCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high
                Cot selected genomic DNA library"
ORIGIN
Query Match      10.2%; Score 18; DB 8; Length 360;
Best Local Similarity 100.0%; Pred. No. 77;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy  84  TAGATCTGTCAGAAACA 101
      |||||
Db    57  TAGATCTGTCAGAAACA 40

RESULT 11
LOCUS      AQ405655/c
DEFINITION HS_5049_B1 G07_SP6E RPCI-11 Human Male BAC Library Homo sapiens
            genomic clone Plate=625 Col=13 Row=N, genomic survey sequence.
ACCESSION  AQ405655
VERSION     AQ405655.1 GI:4415643
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
            1 (bases 1 to 423)
REFERENCE   1 (bases 1 to 423)
AUTHORS     Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
            Kellar,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
            Hood,L.
            Sequence-tagged connectors: A sequence approach to mapping and
            scanning the human genome
            Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
            9380589
            MEDLINE 10449764
            PUBMED
            Contact: Mahairas GG, Wallace JC, Hood L
            High Throughput Sequencing Center
            University of Washington
            401 Queen Anne Avenue North, Seattle, WA 98109, USA

Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Clones are derived from the human BAC library RPCI-11. For BAC
library availability, please contact Pieter de Jong
(pieter@dejong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm)
or from Research Genetics (info@resgen.com). BAC end Web Server:
http://www.htsc.washington.edu
Plate: 625 row: N column: 13
Seq primer: SP6
Class: BAC ends
High quality sequence stop: 423.
FEATURES    source
            Location/Qualifiers
                1..423
                /organism="Homo sapiens"
                /mol_type="genomic DNA"
                /db_xref="taxon:9606"
                /clone="Plate=625 Col=13 Row=N"
                /sex="male"
                /clone_lib="RPCI-11 Human Male BAC Library"
                /notes="Vector: pBACe3.6; Site 1: EcoRI; Site 2: EcoRI;
                Male blood DNA was isolated from one randomly chosen donor
                and partially digested with a combination of EcoRI and
                EcoRI Methyase. Size selected DNA was cloned into the
                pBACe3.6 vector at EcoRI sites"
ORIGIN
Query Match      10.2%; Score 18; DB 8; Length 423;
Best Local Similarity 100.0%; Pred. No. 78;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy  141  GTGTGAATTTCTGAAGGA 158
      |||||
Db    260  GTGTGAATTTCTGAAGGA 243

RESULT 12
LOCUS      CF794539
DEFINITION 890065 MARC 4PIG Sus scrofa CDNA 5', mRNA sequence.
ACCESSION  CF794539
VERSION     CF794539.1 GI:37799112
KEYWORDS    EST.
SOURCE      Sus scrofa (pig)
ORGANISM    Sus scrofa
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
            1 (bases 1 to 466)
REFERENCE   1 (bases 1 to 466)
AUTHORS     Smith,T.P.L., Preking,B.A., Ford,J.J., Vallet,J.L., Wise,T.A.,
            Nonneman,D.J., Wray,J.E. and Keele,J.W.
            Porcine EST collection using a normalized library constructed from
            embryos representing early developmental stages
            Unpublished (2003)
            Contact: Smith TPL
            USDA, ARS, US Meat Animal Research Center
            PO Box 166, Clay Center, NE 68933-0166, USA
            Tel: 402 762 4390
            Fax: 402 762 4390
            Email: smith@email.marc.usda.gov
            Single pass sequencing. Bases called with phred v0.020425.c and
            trimmed with the aid of the trim_alt option. Vector identified with
            cross_match v0.990329.
            Plate: TMW8014 row: E column: 12
            Seq primer: GTAATACGACTCACTATAGG.
FEATURES    source
            Location/Qualifiers
                1..466
                /organism="Sus scrofa"
                /mol_type="mRNA"
                /db_xref="taxon:9823"
                /tissue_type="pooled"
                /lab_host="DH10B"
                /clone_lib="MARC 4PIG"

```

/note="Vector: pCDNA3.1; Site_1: EcoRI; Site_2: NotI;
Library made with combined RNA from day-10, day-13,
day-15, day-25, and day-30 whole embryos."

```

ORIGIN
Query Match      10.2%; Score 18; DB 7; Length 466;
Best Local Similarity 100.0%; Pred. No. 78;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 145 GAATTTCTGAAGGATGAG 162
|||||
DB 108 GAATTTCTGAAGGATGAG 125

RESULT 13
LOCUS AZ644277 466 bp DNA linear GSS 14-DEC-2000
DEFINITION IM0508H1F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
clone UUGC1M0508H1 F, genomic survey sequence.
ACCESSION AZ644277
VERSION AZ644277.1 GI:11772649
KEYWORDS GSS
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 466)
AUTHORS Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C.,
Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T.,
Reilly, M., Rose, M., Rose, R., Stokes, R., Tingey, A., von
Niederhausern, A. and Wright, D., Weiss, R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0508 row: H column: 11
Seq primer: CGTGTGAACGACGCGCCAGT
Class: plasmid ends
High quality sequence stop: 466.
Location/Qualifiers
1..466
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC1M0508H1"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/note="Vector: PWD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adaptored DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of PWD42 (gi:4732114|gb|AF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adaptored mouse DNA was annealed to

```

```

ORIGIN
Query Match      10.2%; Score 18; DB 8; Length 466;
Best Local Similarity 100.0%; Pred. No. 78;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 142 TGTGAATTTCTGAAGGAT 159
|||||
DB 102 TGTGAATTTCTGAAGGAT 119

RESULT 14
LOCUS BG603242/c 477 bp mRNA linear EST 14-AUG-2001
DEFINITION EST502332 Plasmodium yoelii sporozoite cDNA Plasmodium yoelii cDNA
clone PYCDP45, mRNA sequence.
ACCESSION BG603242
VERSION BG603242.1 GI:15153256
KEYWORDS EST
SOURCE Plasmodium yoelii
ORGANISM Plasmodium yoelii
REFERENCE 1 (bases 1 to 477)
AUTHORS Kappe, S.H.I., Gardner, M.J., Brown, S.M., Ross, J., Matuschewski, K.,
Ribeiro, J.M., Adams, J.H., Quackenbush, J., Cho, J., Carucci, D.J.,
Hoffman, S.L. and Nussenzweig, V.
TITLE Exploring the transcriptome of the malaria sporozoite stage
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 98 (17), 9895-9900 (2001)
MEDLINE 21396555
PUBMED 11493695
COMMENT Contact: Malcolm J. Gardner
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301 838 3519
Fax: 301 838 0208
Email: gardner@tigr.org
Request for clones, please contact: Stefan Kappe,
kappes01@popmail.med.nyu.edu Michael Heidelberger Division,
Department of Pathology New York University School of Medicine.
Location/Qualifiers
1..477
/organism="Plasmodium yoelii"
/mol_type="mRNA"
/strain="17XNL"
/db_xref="taxon:5861"
/clone="PYCDP45"
/dev_stage="sporozoites from salivary gland"
/lab_host="E. coli TOP10"
/clone_lib="Plasmodium yoelii sporozoite cDNA"
/note="Vector: pCR4; TA cloning; Plasmodium yoelii
sporozoite cDNA library from salivary gland sporozoites 14
days post-infection"

FEATURES
source
1..477
/organism="Plasmodium yoelii"
/mol_type="mRNA"
/strain="17XNL"
/db_xref="taxon:5861"
/clone="PYCDP45"
/dev_stage="sporozoites from salivary gland"
/lab_host="E. coli TOP10"
/clone_lib="Plasmodium yoelii sporozoite cDNA"
/note="Vector: pCR4; TA cloning; Plasmodium yoelii
sporozoite cDNA library from salivary gland sporozoites 14
days post-infection"

ORIGIN
Query Match      10.2%; Score 18; DB 4; Length 477;
Best Local Similarity 100.0%; Pred. No. 78;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 134 AATCGTGTGTGAATTC 151
|||||
DB 221 AATCGTGTGTGAATTC 204

RESULT 15
LOCUS PT003115U/c 483 bp DNA linear GSS 29-MAY-2003
DEFINITION Paramescium tetraurelia sequence M03E08u of the end of plasmid
PT003115, genomic survey sequence.
ACCESSION AL446451

```

adaptored vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

VERSION AL446451.1 GI:11121726
KEYWORDS GSS.
SOURCE Parametium tetraurelia
ORGANISM Parametium tetraurelia
Eukaryota; Alveolata; Ciliophora; Oligohymenophorea; Peniculida;
Parametium.

REFERENCE 1
AUTHORS Keller, A.M. and Cohen, J.
TITLE An indexed genomic library for Parametium complementation cloning
JOURNAL J. Eukaryot. Microbiol. 47 (1), 1-6 (2000)
MEDLINE 20114709
PUBMED 10651287

REFERENCE 2 (bases 1 to 483)
AUTHORS Dessen, P., Zagulski, M., Gromadka, R., Plattner, H., Kismehl, R.,
Meyer, E., Bernier, M., Schultz, J.E., Linder, J.U., Pearlman, R.E.,
Kung, C., Forney, J., Satir, B.H., Van Houten, J.L., Keller, A.M.,
Froissard, M., Sperling, L. and Cohen, J.
TITLE Parametium genome survey: a pilot project
JOURNAL Trends Genet. 17 (6), 306-308 (2001)
MEDLINE 21273563
PUBMED 11377780

REFERENCE 3 (bases 1 to 483)
AUTHORS Gromadka, R. and Zagulski, M.
TITLE Random sequencing of the Parametium macronuclear genome
JOURNAL Unpublished
REMARK Institute of Biochemistry and Biophysics, Polish Academy of
Sciences, Warsaw, Poland

REFERENCE 4 (bases 1 to 483)
AUTHORS Cohen, J. and Sperling, L.
TITLE Direct Submission
JOURNAL Submitted (01-NOV-2000) Parametium Genome Survey Project, Centre de
Genetique Moleculaire, Centre National de la Recherche
Scientifique, 91198 Gif-sur-Yvette, France. E-mail:
sperling@cgm.cnrs-gif.fr

COMMENT The present survey of the Parametium tetraurelia macronuclear
genome consists of end sequences of a library of random 4-12 kb
fragments obtained by Sau3A partial digestion of macronuclear DNA
cloned in the BamHI site of pBSIIKS-. See [4].
Genes are predicted from matches to other sequences. For more
information about this sequence or the Parametium Project, see
http://parametium.cgm.cnrs-gif.fr.

FEATURES
source
1..483
Location/Qualifiers
/organism="Parametium tetraurelia"
/macronuclear
/mol_type="genomic DNA"
/strain="stock d4-2"
/db_xref="taxon:5888"

ORIGIN
Query Match 10.2%; Score 18; DB 9; Length 483;
Best Local Similarity 100.0%; Pred. No. 79;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 146 AATTTCGAAGGATGAGA 163
|||||
Db 418 AATTTCGAAGGATGAGA 401

Search completed! August 26, 2005, 23:31:23
Job time : 2114.25 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 26, 2005, 18:43:08 ; Search time 324.757 Seconds
(without alignments)
3226.399 Million cell updates/sec

Title: US-10-617-978-14_COPY_64_240

Perfect score: 177

Sequence: 1 gctgacgtcccggaacta.....atgagaacgtgaaggtctga 177

Scoring table: OLIGO NUC

Gapop 60.0 , Gapext 60.0

Searched: 4390206 seqs, 2959870667 residues

Word size : 0

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N Geneseq 16Dec04:*

- 1: Geneseqn1980s:*
- 2: Geneseqn1990s:*
- 3: Geneseqn2000s:*
- 4: Geneseqn2001as:*
- 5: Geneseqn2001bs:*
- 6: Geneseqn2002as:*
- 7: Geneseqn2002bs:*
- 8: Geneseqn2003as:*
- 9: Geneseqn2003bs:*
- 10: Geneseqn2003cs:*
- 11: Geneseqn2003ds:*
- 12: Geneseqn2004as:*
- 13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	19	10.7	143412	11 ACN44512	Acn44512 Mouse gen
2	18	10.2	2365	13 ADQ85216	Adq85216 Human tum
3	18	10.2	3408	10 ADH73022	Adh73022 Human MEG
C 4	18	10.2	3408	10 ADH73024	Adh73024 Human MEG
5	18	10.2	3909	10 ADI60469	Adi60469 Secreted
6	18	10.2	4790	5 AAS45254	Aas45254 cDNA enco
7	18	10.2	5587	6 ADH48775	Adh48775 NOV25 cod
8	18	10.2	5715	10 ADH73025	Adh73025 Human MEG
9	18	10.2	6846	12 ADN04593	Adn04593 Antipsori
10	18	10.2	6846	12 ADQ20312	Adq20312 Human sof
11	18	10.2	7037	5 AAS45066	Aas45066 cDNA enco
12	18	10.2	7132	8 ABX63780	Abx63780 Human cDN
13	18	10.2	7973	10 ADD93418	Add93418 Human lip
14	18	10.2	7974	8 AD47370	Ad47370 Human Lp2
C 15	17	9.6	396	4 AAF94502	Aaf94502 Human ova
C 16	17	9.6	396	6 ABL48852	Abi48852 Ovarian c
C 17	17	9.6	396	6 AFT03169	Aft03169 Human ova
C 18	17	9.6	396	11 ADM10762	Adm10762 Human ova
C 19	17	9.6	396	12 ADJ11092	Adj11092 Represent
C 20	17	9.6	396	12 ADM43353	Adm43353 Human ova

21	17	9.6	468	10 AAD62548	Aad62548 Human T1R
C 22	17	9.6	758	6 ABS77229	Abse77229 Frog embr
C 23	17	9.6	812	3 AAZ89353	Aaz89353 Human 18.
C 24	17	9.6	1014	5 AAH67435	Aah67435 C glutami
C 25	17	9.6	1083	4 AAF71267	Aaf71267 Corynebac
26	17	9.6	1350	8 ACA01824	Aca01824 C. glutam
27	17	9.6	1353	5 AAH65015	Aah65015 C glutami
C 28	17	9.6	1784	10 ADI02464	Adi02464 Human cDN
C 29	17	9.6	2500	4 AAK94785	Aak94785 Human ful
C 30	17	9.6	2500	12 ADL31859	Adl31859 Full leng
C 31	17	9.6	2667	4 ABA06424	Abao6424 Human cDN
C 32	17	9.6	2667	6 ABV83761	Abv83761 Human pol
C 33	17	9.6	2690	4 AAS28803	Aas28803 Human imm
C 34	17	9.6	2690	10 ADB31528	Adb31528 Human cDN
C 35	17	9.6	3024	12 ADQ67269	Adq67269 Novel hum
C 36	17	9.6	3598	2 AAX37725	Aax37725 Human PRO
C 37	17	9.6	3662	2 AAX52264	Aax52264 Protein P
C 38	17	9.6	3662	3 AAZ52206	Aaz52206 Human PRO
C 39	17	9.6	3662	3 ADC78601	Adc78601 Human PRO
C 40	17	9.6	3662	4 AAF72422	Aaf72422 Human PRO
C 41	17	9.6	3662	4 AAS00161	Aas00161 Human cDN
C 42	17	9.6	3662	8 ACA60239	Aca60239 Human cDN
C 43	17	9.6	3662	8 ACD07639	Acd07639 Novel hum
C 44	17	9.6	3662	8 ABX71687	Abx71687 Human cDN
C 45	17	9.6	3662	8 ACH07019	Ach07019 Human sec

ALIGNMENTS

RESULT 1

ACN44512/c
ID ACN44512 standard; DNA; 143412 BP.

XX AC ACN44512;

DT 18-NOV-2004 (first entry)

XX DB Mouse genomic sequence MCG20543.

XX KW Cytostatic; carcinoma; lymphoma; cancer; murine; gene; ss.

XX OS Mus musculus.

XX PN WO2003073826-A2.

XX PD 12-SEP-2003.

XX PF 28-FEB-2003; 2003WO-US006235.

XX PR 01-MAR-2002; 2002US-00087192.

XX PA (SAGR-) SAGRES DISCOVERY.

XX PI Morris DW;

XX DR WPI; 2003-328604/31.

XX PT Recombinant nucleic acid useful for diagnosis and treatment of carcinoma comprises a nucleotide sequence.

XX PS Claim 1; SEQ ID NO 997; Opp; English.

CC The present invention relates to novel DNA and protein sequences which are associated with carcinomas. The sequences are useful for: (i) for screening drug candidates; (ii) for screening of bioactive agent capable of binding to Carcinoma Associated Protein (CAP); (iii) for screening of a bioactive agent capable of modulating the activity of CAP; (iv) for evaluating the effect of a candidate carcinoma drug; (v) for diagnosing carcinoma; (vi) for inhibiting the activity of CAP; (vii) for treating carcinoma; (viii) for neutralizing the effect of CAP; (ix) as a biochip; (x) for diagnosing carcinoma or a propensity to carcinoma; and (xi) for determining Carcinoma Associated (CA) gene copy number. In addition, the

CA genes are useful as DNA vaccines and the CAP are useful as markers of carcinoma including lymphoma. The present sequence is one such CA coding sequence. Note: This patent is an equivalent to basic patent US2002192586A1, for which no sequence data was published

```
Query Match      10.7%; Score 19; DB 11; Length 143412;
Best Local Similarity 100.0%; Pred. No. 3.5;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 127 GCCTTCCAATGCTGGTGTG 145

Db 129843 GCCTTCCAAATGCTGGTGTG 129825

RESULT 2
ADQ85216
ID ADQ85216 standard; cDNA; 2365 BP.

07-OCT-2004 (first entry)

DE Human tumour-associated antigenic target (TAT) cDNA sequence #2030.

human; tumour-associated antigenic target; TAT; cytostatic; gene therapy;
KW cancer; cell proliferative disorder; gene; ss.

OS Homo sapiens.

PN WO2004060270-A2.

22-JUL-2004,

15-OCT-2003, 2003WO-US029126.

18-OCT-2002: 2002US-0418988P.

PA (GETH) GENENTECH INC.

PA (WUTD/) WU T D.

PA (ZHOU/) ZHOU Y.

PI Wu TD, Zhou Y;

DR WPI; 2004-534300/51.

PT New nucleic acid molecule and encoded polypeptide, for diagnosing,
PT preventing or treating cell proliferative disorders such as cancer.

PS Claim 1: SEQ ID NO 2030: 5504pp: English.

The present invention describes an isolated tumour-associated antigenic target (TAT) nucleic acid comprising: (a) any of 4622 nucleotide sequences (see SEQ ID NO:1 to 4622); (b) the full-length coding region of (a); (c) the complement of (a) or (b); (d) a sequence that has 80% sequence identity to (a) - (c); or (e) a sequence that hybridises to (a) - (c). Also described: (1) an expression vector comprising the above nucleic acid; (2) a host cell comprising the above expression vector; (3) a process for producing a polypeptide; (4) an isolated polypeptide comprising: (a) an amino acid sequence encoded by any of the above nucleotide sequences; (b) an amino acid sequence encoded by the full-length coding region of the above nucleotide sequences; or (c) a sequence having at least 80% identical to (a) or (b); (5) a chimeric polypeptide comprising the above polypeptide fused to a heterologous polypeptide; (6) an isolated antibody that binds to the above polypeptide; (7) a process for producing the antibody; (8) an isolated oligopeptide that binds to the above polypeptide; (9) a tumour-associated antigenic target (TAT) binding organic molecule that binds to the above polypeptide; (10) a composition of matter comprising the above (chimeric) polypeptide, antibody, oligopeptide or TAT binding organic molecule, in combination with a carrier; (11) an article of manufacture comprising a container and the composition of matter contained within the container; (12) methods of

PS Disclosure; SEQ ID NO 1; 47pp; English.

XX This invention relates to a novel human MEGF7 polypeptide (a polypeptide
CC having multiple epidermal growth factor-like domains, including an
CC unidentified low density lipoprotein (LDL) receptor-like protein). The
CC sequences of the invention may be useful for gene therapy or protein
CC therapy or for the modulation of MEGF7 expression and activity. The
CC invention may be useful for the development of methods to diagnose or
CC treat diseases associated with inappropriate MEGF7 activity or levels.
CC The polypeptides and polynucleotides are also useful for configuring
CC screening methods for detecting the effect of added compounds on the
CC production of mRNA and polypeptide in cells. The polynucleotides are also
CC useful as diagnostic reagents, through detecting mutations in the
CC associated gene. The polynucleotide sequence is useful for chromosome
CC localisation studies and tissue expression studies. The present sequence
CC is that of a cDNA sequence which is related to the invention.

XX
SQ Sequence 3408 BP; 752 A; 943 C; 1028 G; 685 T; 0 U; 0 Other;

Query Match 10.2%; Score 18; DB 10; Length 3408;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 130 TTCCAATGCTGGTGTGAA 147
|||||

DB 1411 TTCCAATGCTGGTGTGAA 1428
|||||

RESULT 4

ADH73024/c

ID ADH73024 standard; DNA; 3408 BP.

XX
AC ADH73024;

XX
DT 25-MAR-2004 (first entry)

XX
DE Human MEGF7-related DNA sequence SeqID3.

XX
KW MEGF7; epidermal growth factor-like domain; low density lipoprotein;
KW LDL receptor-like protein; gene therapy; protein therapy;
KW MEGF7 expression; ds; human.

XX
OS Homo sapiens.

XX
PN GB2381790-A.

XX
PD 14-MAY-2003.

XX
PF 26-SEP-2002; 2002GB-00022372.

XX
PR 26-SEP-2001; 2001GB-00023124.

XX
PR 26-JUN-2002; 2002GB-00014703.

XX
PA (GLAXO) GLAXO GROUP LTD.

XX
PI Volpe F;

XX
DR WPI; 2003-432835/41.

XX
PT Novel isolated human MEGF7 polypeptide, a polypeptide having multiple
PT epidermal growth factor-like domains, useful in diagnostic assays for
PT detecting diseases associated with inappropriate MEGF7 activity or
PT levels.

XX
PS Disclosure; SEQ ID NO 3; 47pp; English.

XX This invention relates to a novel human MEGF7 polypeptide (a polypeptide
CC having multiple epidermal growth factor-like domains, including an
CC unidentified low density lipoprotein (LDL) receptor-like protein). The
CC sequences of the invention may be useful for gene therapy or protein
CC therapy or for the modulation of MEGF7 expression and activity. The
CC invention may be useful for the development of methods to diagnose or
CC treat diseases associated with inappropriate MEGF7 activity or levels.

XX
SQ Sequence 3408 BP; 752 A; 943 C; 1028 G; 685 T; 0 U; 0 Other;

Query Match 10.2%; Score 18; DB 10; Length 3408;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 130 TTCCAATGCTGGTGTGAA 147
|||||

DB 1998 TTCCAATGCTGGTGTGAA 1981
|||||

RESULT 5

AD160469

ID AD160469 standard; DNA; 3909 BP.

XX
AC AD160469;

XX
DT 15-APR-2004 (first entry)

XX
DE Secreted polypeptide encoding gene #8.

XX
KW ds; gene; osteopathic; vulnery; cytostatic; gene therapy; diagnosis;
KW forensics; gene mapping; mutation identification; biodiversity;
KW chromosome marker; immune response; myeloid cell disorder;
KW lymphoid cell disorder; bone cartilage; tendon; ligament;
KW nerve tissue growth; wound healing; burns; incision; ulcer; cancer.

XX
OS Homo sapiens.

XX
PN WO2003025142-A2.

XX
PD 27-MAR-2003.

XX
PF 18-SEP-2002; 2002WO-US029636.

XX
PR 18-SEP-2001; 2001US-0323349P.

XX
PR 16-SEP-2002; 2002US-00323349.

XX
PA (HYSE-) HYSEQ INC.

XX
PI Tang YT, Asundi V, Goodrich RW, Ren F, Zhang J, Zhao QA, Wang J;
PI Ghosh M, Xue AJ, Wehrman T, Weng G, Zhou P, Drmanac RT;

XX
DR WPI; 2003-354601/33.

XX
DR P-PSDB; AD160124.

XX
PT New polynucleotides and secreted proteins, useful for treating myeloid or
PT lymphoid cell disorders, in bone cartilage, tendon, ligament and nerve
PT tissue growth or regeneration, in wound healing, and in tissue repair and
PT replacement.

XX
PS Claim 1; SEQ ID NO 8; 243pp; English.

XX The invention relates to novel isolated polynucleotides or a sequence
CC encoding a polypeptide with biological activity, where the polynucleotide
CC hybridizes to the polynucleotide under stringent hybridization conditions
CC or has greater than 99% sequence identity with the polynucleotide. The
CC polynucleotides and polypeptides are useful in diagnostics, forensics,
CC gene mapping, identification of mutations responsible for genetic
CC disorders and other traits, to assess biodiversity, as nutritional
CC sources or supplements. The polynucleotides may also be used as molecular
CC weight markers, chromosome markers or map related gene positions, or as
CC an antigen to raise anti-DNA antibodies or elicit immune response. The
CC polypeptides are useful for raising antibodies, as markers for tissues in
CC which the corresponding polypeptide is expressed, for re-engineering

CC damaged or diseased tissues, for treating myeloid or lymphoid cell
 CC disorders, in bone cartilage, tendon, ligament and/or nerve tissue growth
 CC or regeneration, in wound healing, in tissue repair and replacement, in
 CC healing of burns, incisions and ulcers, and in treating cancer. This
 CC sequence corresponds to a polynucleotide sequence of the invention.

XX Sequence 3909 BP; 836 A; 1092 C; 1168 G; 823 T; 0 U; 0 Other;

Query Match 10.2%; Score 18; DB 10; Length 3909;
 Best Local Similarity 100.0%; Pred. No. 13;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 130 TTCCAATGCTGCTGTGAA 147

Db 1507 TTCCAATGCTGCTGTGAA 1524

RESULT 6

AAS45254

ID AAS45254 standard; cDNA; 4790 BP.

XX AAS45254;

DT 18-DEC-2001 (first entry)

DE cDNA encoding novel human secretory protein, Seq ID No 523.

XX Human; secreted protein; arthritis; Crohn's disease; sepsis; shock;
 KW ischaemia-reperfusion injury; haematopoiesis; cancer; neuropathy;
 KW transgenic animal; Alzheimer's disease; Parkinson's disease; burn;
 KW amyotrophic lateral sclerosis; platelet disorder; thrombocytopenia;
 KW ulcer; osteoporosis; bone degenerative disorder; periodontal disease;
 KW gut protection; lung; liver fibrosis; immune deficiency; infection;
 KW severe combined immunodeficiency; SCID; autoimmune disorder; allergy;
 KW multiple sclerosis; rheumatoid arthritis; diabetes mellitus; asthma;
 KW fertility; analgesic; pain; antigen; ss.

OS Homo sapiens.

FN WO200166889-A2.

XX 13-SEP-2001.

XX 05-MAR-2001; 2001WO-US004942.

XX 07-MAR-2000; 2000US-00519705.

PR 19-MAY-2000; 2000US-00574454.

PR 17-JUN-2000; 2000US-00596193.

PR 14-JUL-2000; 2000US-00616847.

PR 19-SEP-2000; 2000US-00656363.

PR 20-OCT-2000; 2000US-00693367.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Auandi V, Xu C, Wehrman T, Ren P, Ma Y, Zhou P;

PI Zhao QA, Yang Y, Drmanac RT, Zhang J, Chen R, Xue AJ, Wang J;

XX WPI; 2001-589934/66.

DR P-PSDB; AAU28354.

XX Novel polypeptides and nucleic acids obtained from cDNA libraries

PT prepared from various human tissues, for diagnosis and treatment of

PT cancer, neurological, inflammatory, and autoimmune disorders.

XX Claim 1; SEQ ID NO 523; 107pp; English.

XX The invention relates to novel isolated human secreted polypeptides (I)

CC and polynucleotides (II). (I) and (II) are useful for treating

CC inflammatory conditions such as arthritis, nephritis, Crohn's disease,

CC creating transgenic animals useful for studying the in vivo activities of
 CC the polypeptide as well as for studying modulators of the polypeptides.

CC (I) induces the proliferation of neural cells and regeneration of nerve
 CC and brain tissue and is useful for the treatment of central and

CC peripheral nervous system diseases and neuropathies, such as Alzheimer's,
 CC Parkinson's disease, Huntington's disease, and amyotrophic lateral

CC sclerosis. In addition, (I) is involved in chemotactic or chemokinetic
 CC activity, regulation of haematopoiesis and is useful for treating myeloid

CC or lymphoid cell disorders, platelet disorders such as thrombocytopenia
 CC and for regeneration of bone, cartilage, tendon, ligament and/or nerve

CC tissue growth, and in tissue repair, healing of burns, incisions, ulcers,
 CC for treating osteoporosis, osteoarthritis, bone degenerative disorders,

CC or periodontal disease. Furthermore, (I) is also useful for gut
 CC protection or regeneration and treatment of lung or liver fibrosis,

CC reperfusion injury in various tissues, various immune deficiencies and
 CC disorders including severe combined immunodeficiency (SCID), bacterial or

CC fungal infections, autoimmune disorders e.g. multiple sclerosis,
 CC rheumatoid arthritis, diabetes mellitus, myasthenia gravis, allergic

CC reactions and conditions, such as asthma or other respiratory problems.
 CC In addition, (I) affects biorhythms or circadian cycles of rhythms,

CC fertility, metabolism, catabolism, anabolism, storage or elimination of
 CC dietary fat, lipid, protein, carbohydrate, vitamins, minerals, provides

CC analgesic effects or other pain reducing effects, immunoglobulin like
 CC activity and can act as an antigen in a vaccine composition to raise an

CC immune response. AAS44920-AAS45295 represent novel human secreted protein
 CC coding sequences of the invention

XX Sequence 4790 BP; 1126 A; 1336 C; 1368 G; 960 T; 0 U; 0 Other;

Query Match 10.2%; Score 18; DB 5; Length 4790;

Best Local Similarity 100.0%; Pred. No. 13;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 130 TTCCAATGCTGCTGTGAA 147

Db 321 TTCCAATGCTGCTGTGAA 338

RESULT 7

ADH48775

ID ADH48775 standard; DNA; 5587 BP.

XX ADH48775;

XX 25-MAR-2004 (first entry)

DT NOV25 coding sequence, SEQ ID 59.

XX Human; NOVX; atherosclerosis; hypertension; obesity; cancer; cytostatic;

XX hypotensive; antiarteriosclerotic; anorectic; gene therapy; NOV25;

XX MEGF7-like protein; chromosome 11; gene; ds; SNP;

XX single nucleotide polymorphism.

XX Homo sapiens.

XX Key Location/Qualifiers

FT Variation replace(1836,T)

FT FT /*tag= a

FT /standard_name= "Single nucleotide polymorphism"

XX WO200268652-A2.

XX 06-SEP-2002.

XX 26-FEB-2002; 2002WO-US005910.

XX 26-FEB-2001; 2001US-0271646P.

PR 27-FEB-2001; 2001US-0271840P.

PR 28-FEB-2001; 2001US-0272404P.

PR 28-FEB-2001; 2001US-0272405P.

PR 28-FEB-2001; 2001US-0272410P.

PR 28-FEB-2001; 2001US-0272414P.

PR 02-MAR-2001; 2001US-0272787P.

```
PR 02-MAR-2001; 2001US-0272922P.
PR 02-MAR-2001; 2001US-0273048P.
PR 02-MAR-2001; 2001US-0273300P.
PR 16-MAR-2001; 2001US-0276401P.
PR 20-MAR-2001; 2001US-0277324P.
PR 20-MAR-2001; 2001US-0278660P.
PR 30-MAR-2001; 2001US-0280039P.
PR 30-MAR-2001; 2001US-0280234P.
PR 02-APR-2001; 2001US-0280818P.
PR 12-APR-2001; 2001US-0283443P.
PR 23-APR-2001; 2001US-0285754P.
PR 24-APR-2001; 2001US-0286098P.
PR 03-MAY-2001; 2001US-0288353P.
PR 17-MAY-2001; 2001US-0291703P.
PR 31-MAY-2001; 2001US-0294834P.
PR 20-JUN-2001; 2001US-0295695P.
PR 21-JUN-2001; 2001US-0299845P.
PR 05-JUL-2001; 2001US-0303242P.
PR 13-AUG-2001; 2001US-0311981P.
PR 16-AUG-2001; 2001US-0312858P.
PR 17-AUG-2001; 2001US-0313280P.
PR 29-AUG-2001; 2001US-0315614P.
PR 17-SEP-2001; 2001US-0322818P.
PR 25-FEB-2002; 2002US-00322818.
XX
PA (CURA-) CURAGEN CORP.
XX
XX Alsbrook JP, Anderson DW, Ballinger RA, Boldog FL, Burgess CE;
PI Casman SJ, Ellerman KE, Gangolli EA, Gerlach VL, Gilbert JA;
PI Gorman L, Guo X, Gusev VT, Kekuda R, Li L, Liu X, Malyankar UM;
PI Miller CE, Millet I, Padigar M, Patturajan M, Pena CEA, Peyman JA;
PI Rastelli L, Shenoy SG, Shinkets RA, Smithson G, Spytek KA, Stone DJ;
PI Taupier RJ, Tchernev VT, Vernet CAM, Zerhusen BD;
XX
XX WPI; 2002-698672/75.
DR P-PSDB; ADH48776.
XX
XX New NOVX polypeptides or polynucleotides, useful for preventing or
PT treating disorders or syndromes e.g., atherosclerosis, hypertension,
PT obesity or cancer.
XX
XX Claim 8; Page 156-157; 923pp; English.
XX
XX The present invention relates to novel human NOVX proteins, where X is
CC any number from 1 to 91 and their coding sequences. The proteins and
CC coding sequences are useful for preventing or treating disorders or
CC syndromes e.g. atherosclerosis, hypertension, obesity or cancer. NOV25 is
CC a MEGF7-like protein and its coding sequence maps to chromosome 11.
XX
XX Sequence 5587 BP; 1297 A; 1556 C; 1610 G; 1124 T; 0 U; 0 Other;
SQ
Query Match 10.2%; Score 18; DB 6; Length 5587;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 130 TTCCAATGCTGGTGTGAA 147
| | | | | | | | | | | | | | | | | |
DB 1087 TTCCAATGCTGGTGTGAA 1104
RESULT 8
ADH73025
ID ADH73025 standard; DNA; 5715 BP.
XX
XX ADH73025;
AC
XX
XX 25-MAR-2004 (first entry)
DT
XX
XX Human MEGF7 gene sequence.
DE
XX
XX MEGF7; epidermal growth factor-like domain; low density lipoprotein;
KW LDL receptor-like protein; gene therapy; protein therapy;
KW MEGF7 expression; human; ds; gene.
XX
```

```
XX Homo sapiens.
OS
XX
XX Key Location/Qualifiers
XX CDS 1..5715
FT /*tag= a
FT /product= "Human MEGF7 protein"
FT /partial
FT /note= "No stop codon"
XX
XX GB2381790-A.
PN
XX
XX 14-MAY-2003.
PD
XX
XX 26-SEP-2002; 2002GB-00022372.
PF
XX
XX 26-SEP-2001; 2001GB-00023124.
PR
XX 26-JUN-2002; 2002GB-00014703.
PR (GLAX ) GLAXO GROUP LTD.
PA
XX
XX Volpe F;
PI
XX
XX WPI; 2003-432835/41.
DR P-PSDB; ADH73026.
XX
XX Novel isolated human MEGF7 polypeptide, a polypeptide having multiple
PT epidermal growth factor-like domains, useful in diagnostic assays for
PT detecting diseases associated with inappropriate MEGF7 activity or
PT levels.
XX
XX Claim 1; SEQ ID NO 4; 47pp; English.
XX
XX This invention relates to a novel human MEGF7 polypeptide (a polypeptide
CC having multiple epidermal growth factor-like domains, including an
CC unidentified low density lipoprotein (LDL) receptor-like protein). The
CC sequences of the invention may be useful for gene therapy or protein
CC therapy or for the modulation of MEGF7 expression and activity. The
CC invention may be useful for the development of methods to diagnose or
CC treat diseases associated with inappropriate MEGF7 activity or levels.
CC The polypeptides and polynucleotides are also useful for configuring
CC screening methods for detecting the effect of added compounds on the
CC production of mRNA and polypeptide in cells. The polynucleotides are also
CC useful as diagnostic reagents, through detecting mutations in the
CC associated gene. The polynucleotide sequence is useful for chromosome
CC localisation studies and tissue expression studies. The present sequence
CC is the gene which encodes the human MEGF7 protein of the invention.
XX
XX Sequence 5715 BP; 1316 A; 1588 C; 1665 G; 1146 T; 0 U; 0 Other;
SQ
Query Match 10.2%; Score 18; DB 10; Length 5715;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 130 TTCCAATGCTGGTGTGAA 147
| | | | | | | | | | | | | | | | | |
DB 1246 TTCCAATGCTGGTGTGAA 1263
RESULT 9
ADN04593
ID ADN04593 standard; cDNA; 6846 BP.
XX
XX
XX AC ADN04593;
XX
XX 01-JUL-2004 (first entry)
DT
XX
XX Antipsoriatic cDNA sequence #504.
DE
XX
XX ds; gene; antipsoriatic; gene therapy; psoriasis; diagnosis.
KW
XX
XX Homo sapiens.
OS
XX
```


XX PS Claim 1; SEQ ID NO 147; 107pp; English.

XX CC The invention relates to novel isolated human secreted polypeptides (I)

XX CC and polynucleotides (II). (I) and (II) are useful for treating

XX CC inflammatory conditions such as arthritis, nephritis, Crohn's disease,

XX CC ischaemia-reperfusion injury, shock, sepsis, immune responses, and is

XX CC involved in increasing haematopoiesis, stem cell survival, bone growth

XX CC and remodeling. (II) and modulators of (II) are useful for

XX CC prophylaxis or treatment of one or more cancers. (II) is also useful for

XX CC creating transgenic animals useful for studying the in vivo activities of

XX CC the polypeptide as well as for studying modulators of the polypeptides.

XX CC (I) induces the proliferation of neural cells and regeneration of nerve

XX CC and brain tissue and is useful for the treatment of Central and

XX CC peripheral nervous system diseases and neuropathies, such as Alzheimer's,

XX CC Parkinson's disease, Huntington's disease, and amyotrophic lateral

XX CC sclerosis. In addition, (I) is involved in chemotactic or chemokinetic

XX CC activity, regulation of haematopoiesis and is useful for treating myeloid

XX CC and lymphoid cell disorders, platelet disorders such as thrombocytopenia

XX CC and for regeneration of bone, cartilage, tendon, ligament and/or nerve

XX CC tissue growth, and in tissue repair, healing of burns, incisions, ulcers,

XX CC for treating osteoporosis, osteoarthritis, bone degenerative disorders,

XX CC or periodontal disease. Furthermore, (I) is also useful for gut

XX CC protection or regeneration and treatment of lung or liver fibrosis,

XX CC reperfusion injury in various tissues, various immune deficiencies and

XX CC disorders including severe combined immunodeficiency (SCID), bacterial or

XX CC fungal infections, autoimmune disorders e.g. multiple sclerosis,

XX CC rheumatoid arthritis, diabetes mellitus, myasthenia gravis, allergic

XX CC reactions and conditions, such as asthma or other respiratory problems.

XX CC In addition, (I) affects biorhythms or circadian cycles of rhythms,

XX CC fertility, metabolism, catabolism, anabolism, storage or elimination of

XX CC dietary fat, lipid, protein, carbohydrate, vitamins, minerals, provides

XX CC analgesic effects or other pain reducing effects, immunoglobulin like

XX CC activity and can act as an antigen in a vaccine composition to raise an

XX CC immune response. AA544920-AA545295 represent novel human secreted protein

XX CC coding sequences of the invention

XX SQ Sequence 7037 BP; 1681 A; 1867 C; 1925 G; 1564 T; 0 U; 0 Other;

Query Match 10.2%; Score 18; DB 5; Length 7037;

Best Local Similarity 100.0%; Pred. No. 13;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 130 TTCCAATGCTGGTGTGAA 147

Db 450 TTCCAATGCTGGTGTGAA 467

RESULT 12

ABX63780

ID ABX63780 standard; cDNA; 7132 BP.

XX AC ABX63780;

XX DT 26-FEB-2003 (first entry)

XX DE Human cDNA #780 differentially expressed in activated vascular tissue.

XX KW Human; gene; ss; vascular tissue; cytostatic; atherosclerosis; cardiant;

XX KW hypotensive; antidiabetic; gynaecological; vasotrophic; cerebroprotective;

XX KW gene therapy; vascular disease; cancer; coronary; artery disease;

XX KW hypertension; diabetes; pre-eclampsia; restenosis;

XX KW ischaemia-reperfusion injury; stroke.

XX OS Homo sapiens.

XX FN US2002137081-A1.

XX PD 26-SEP-2002.

XX PF 08-JAN-2002; 2002US-00044090.

XX PR 28-JUL-2000; 2000US-0222469P.

PR 08-JAN-2001; 2001US-0260483P.

XX PA (BAND/) BANDMAN O.

XX PI Bandman O;

XX DR WPI; 2003-110597/10.

XX PT Combination for diagnosing, staging, treating, or monitoring the

XX PT progression of treatment of a vascular disease, e.g. atherosclerosis,

XX PT comprises several cDNAs that are differentially expressed in activated

XX PT vascular tissue.

XX PS Claim 1; Page; 18pp; English.

XX CC This invention relates to a combination comprising several cDNAs that are

XX CC differentially expressed in activated vascular tissue. The invention also

XX CC discloses a high throughput method for detecting differentially expressed

XX CC cDNAs in a sample. The cDNAs of the invention may have

XX CC antiarteriosclerotic; cytostatic; cardiant; hypotensive; antidiabetic;

XX CC gynaecological; vasotrophic and cerebroprotective activities and may be

XX CC used in gene therapy. The cDNAs of the invention may be used in a high-

XX CC throughput methods for detecting differential expression of one or more

XX CC cDNAs in a sample, or screening several molecules or compounds to

XX CC identify a molecule or compound that specifically binds a cDNA of the

XX CC invention. A protein encoded by the cDNA may be used to screen several

XX CC molecules or compounds to identify a ligand that specifically binds to

XX CC the protein, or to produce or purify an antibody to the protein that can

XX CC be used to detect a protein in a sample or purify a natural or

XX CC recombinant protein from a sample. The nucleotides may be useful for

XX CC diagnosing, staging, treating, or monitoring the progression of treatment

XX CC of a vascular disease, e.g. atherosclerosis, cancer, coronary artery

XX CC disease, hypertension, diabetes, pre-eclampsia, ischaemia-reperfusion

XX CC injury, restenosis, or stroke. The cDNAs can also be used for large-scale

XX CC genetic or gene expression analysis of several new nucleic acid

XX CC molecules. Antibodies to the proteins encoded by the cDNAs are useful for

XX CC diagnosing pre-pathologic disorders, and chronic or acute diseases

XX CC associated with abnormalities in the expression, amount or distribution

XX CC of the protein. The present sequence represents a cDNA of the invention

XX CC that is differentially expressed in activated vascular tissue. Note: The

XX CC sequence data for this patent did not form part of the specification, but

XX CC was obtained in electronic format directly from USPTO at

XX CC <http://seqdata.uspto.gov/sequence.html?DocID=20020137081>

XX SQ Sequence 7132 BP; 1702 A; 1879 C; 1942 G; 1609 T; 0 U; 0 Other;

Query Match 10.2%; Score 18; DB 8; Length 7132;

Best Local Similarity 100.0%; Pred. No. 13;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 130 TTCCAATGCTGGTGTGAA 147

Db 259 TTCCAATGCTGGTGTGAA 276

RESULT 13

ADD93418

ID ADD93418 standard; cDNA; 7973 BP.

XX AC ADD93418;

XX DT 29-JAN-2004 (first entry)

XX DE Human lipid-associated molecule LIPAM-6 polynucleotide.

XX KW Human; lipid-associated molecule; LIPAM-6; neuroprotective; relaxant;

XX KW antithyroid; antidiabetic; cytostatic; dermatological; immunosuppressive;

XX KW antiinflammatory; thyromimetic; antiallergic; cerebroprotective;

XX KW gastrointestinal; hepatotropic; nephrotropic; anticonvulsant;

XX KW antiparkinsonian; antibacterial; antiparasitic; fungicide; protozoacide;

XX KW virucide; uropathic; antirheumatic; cardiant; cardiovascular; anti-HIV;

XX KW nootropic; gene; ss.


```
OS XX Homo sapiens.
FH XX
FT XX Key Location/Qualifiers
FT CDS 138..5858
FT sig_peptide /*tag= b
FT FT 138..197
FT FT /*tag= a
FT mat_peptide /product= "Human LIPAM-6"
FT 198..5855
FT /*tag= c
PN WO2003083081-A2.
XX
XX 09-OCT-2003.
XX
XX 27-MAR-2003; 2003WO-US009755.
XX
XX 29-MAR-2002; 2002US-0368722P.
XX 03-MAY-2002; 2002US-0377576P.
XX 05-JUL-2002; 2002US-0393934P.
XX 27-SEP-2002; 2002US-0414269P.
XX
XX (INCY-) INCYTE CORP.
XX
XX Emerling BM; Marquis JP, Chawla NK, Lee SY, Duggan BM, Warren BA;
XX Baughn MR, Lee EA, Griffin JA, Kable AE, Elliott VS, Chang H;
XX Lee S, Ramkumar J, Bulloch SA, Hafalia AJA, Khare R, Jiang X;
XX Jackson AA;
XX
XX WPI; 2003-788347/74.
XX P-PSDB; ADD93399.
XX
XX New LIPAM polypeptides, useful for diagnosing, preventing, and treating
XX disorders associated with abnormal expression or activity of LIPAM, e.g.
XX neuromuscular, immunological, cardiovascular disorders, cancer and/or
XX infections.
XX
XX Claim 80; Page 222-224; 238pp; English.
XX
XX The present sequence is the nucleotide sequence of human lipid-associated
XX molecule LIPAM-6 (incyte polynucleotide 7510885CBI), which encodes a
XX protein that shows sequence homology to chicken alpha-2-macroglobulin
XX receptor. This is one of 19 LIPAM polynucleotides of the invention. The
XX invention relates to novel LIPAMs and the nucleic acids encoding them,
XX and to the use of nucleic acids and proteins in the diagnosis, treatment
XX and prevention of disorders associated with abnormal expression or
XX activity of LIPAM such as neurodegenerative disorders (e.g. Parkinson's
XX disease, Alzheimer's disease), muscular disorders (e.g. myotonic
XX dystrophy, catatonias), endocrine disorders (e.g. diabetes, Grave's
XX disease), cancers (e.g. leukaemia, cervical or breast cancers),
XX immunological disorders (e.g. scleroderma, systemic lupus erythematosus,
XX allergies), gastrointestinal disorders (e.g. Crohn's disease), renal
XX disorders (e.g. Goodpasture's syndrome), infections (e.g. viral,
XX bacterial, fungal, parasitic, protozoal, helminthic), cardiovascular
XX disorders (e.g. atherosclerosis), or hepatic diseases (e.g. cirrhosis).
XX The invention also relates to the assessment of the effects of exogenous
XX compounds on the expression of nucleic acids and LIPAMs. The invention
XX provides expression vectors, host cells, antibodies, agonists and
XX antagonists, transgenic organisms, and arrays and microarrays of the
XX polynucleotides.
XX
XX Sequence 7973 BP; 1849 A; 2140 C; 2244 G; 1740 T; 0 U; 0 Other;
SQ
Query Match 10.2%; Score 18; DB 10; Length 7973;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 130 TTCCAATGCTGGTGA 147
|||||
Db 1386 TTCCAATGCTGGTGA 1403
|||||
RESULT 14
```


CC Parkinson's disease or Alzheimer's disease) or developmental disorders
CC (e.g. Down's syndrome or cerebral palsy). They are also used in gene
CC therapy. The present sequence is human LP288 DNA
XX
SQ Sequence 7974 BP; 1848 A; 2143 C; 2243 G; 1740 T; 0 U; 0 Other;
Query Match 10.2%; Score 18; DB 8; Length 7974;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 130 TTCCAATGCTGCTGTGAA 147
Db 1387 TTCCAATGCTGCTGTGAA 1404
RESULT 15
AAF94902/C
ID AAF94902 standard; cDNA; 396 BP.
XX
AC AAF94902;
XX
XX 23-MAY-2001 (first entry)
XX
DE Human ovarian cancer associated coding sequence SEQ ID NO: 93.
XX
KW Human, ovarian cancer; vaccine; gene therapy; carcinoma; ss.
OS Homo sapiens.
XX
FN WO200118046-A2.
XX
PD 15-MAR-2001.
XX
XX 08-SEP-2000; 2000WO-US024827.
XX
XX 10-SEP-1999; 99US-00394374.
PR 01-MAY-2000; 2000US-00561778.
PR 15-AUG-2000; 2000US-00640173.
PR 07-SEP-2000; 2000US-00656668.
XX
XX (CORI-) CORIXA CORP.
PA
XX
XX Xu J, Stolk JA;
XX
XX WPI; 2001-211395/21.
XX
PT Isolated polypeptides associated with ovarian carcinomas, and the nucleic
PT acids that encode them, useful for the prevention diagnosis and treatment
PT of ovarian cancers.
XX
PS Claim 5; Page 146; 189pp; English.
XX
XX The present invention provides a number of coding sequences and proteins,
CC the over-expression of which is associated with ovarian carcinoma/cancer.
CC These can be used in the diagnosis, treatment and prevention of ovarian
CC cancer, optionally by gene therapy or in the form of a vaccine. The
CC present sequence is an example of one of these sequences
XX
SQ Sequence 396 BP; 110 A; 86 C; 79 G; 112 T; 0 U; 9 Other;
Query Match 9.6%; Score 17; DB 4; Length 396;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 149 TTCTGAGGATGAGAAC 165
Db 199 TTCTGAGGATGAGAAC 183

Search completed: August 26, 2005, 21:18:53
Job time : 326.757 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 26, 2005, 20:14:05 ; Search time 2006.75 Seconds
(without alignments)
3186.647 Million cell updates/sec

Title: US-10-617-978-14_COPY_73_240
Perfect score: 168
Sequence: 1 ccggaaactaccacttg.....atgagaactggaagctctga 168

Scoring table: OLIGO_NUC
Gapop_60.0 , Gapext 60.0

Searched: 34239544 seqs, 19032134700 residues

Word size : 0

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : EST:
1: gb_est1:
2: gb_est2:
3: gb_hc:
4: gb_est3:
5: gb_est4:
6: gb_est5:
7: gb_est6:
8: gb_gss1:
9: gb_gss2:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	22	13.1	435	2 BE950396	UI-M-CEO-
C 2	21	12.5	379	7 CK817241	UMC-bov. 0
C 3	21	12.5	665	1 AJ814218	AJ814218
C 4	21	12.5	674	6 CB166163	KXE603014
C 5	21	12.5	728	7 CK949817	4074871 B
C 6	19	11.3	580	2 AW361823	PMO-CT026
C 7	19	11.3	1101	8 B10170	F14D20-Sp6
C 8	18	10.7	263	2 BS586775	BS586775
C 9	18	10.7	313	4 BG603087	EST502177
C 10	18	10.7	360	8 BZ686719	PUBBU547D
C 11	18	10.7	423	8 AQ405655	HS_5049 B
C 12	18	10.7	466	7 CF794539	890065 MA
C 13	18	10.7	466	8 AZ644277	1M0508H11
C 14	18	10.7	477	4 BG603242	EST502332
C 15	18	10.7	483	9 PT003115U	AL446451
C 16	18	10.7	547	8 BZ675947	PARAMECIU
C 17	18	10.7	554	5 BP615364	BP615364
C 18	18	10.7	587	8 CC900317	ZMBBb023
C 19	18	10.7	589	8 AQ350422	RPC11-11
C 20	18	10.7	610	4 BG603086	EST502176
C 21	18	10.7	611	7 CK898930	SGP162516
C 22	18	10.7	621	7 CF793169	886193 MA
C 23	18	10.7	623	6 CB963735	AGENCOURT
C 24	18	10.7	628	4 BJ623886	BJ623886

25	18	10.7	640	8	AZ710307	AZ710307
26	18	10.7	656	2	BF611918	RPC1-24-1
27	18	10.7	675	5	EX850657	de89h10.Y
28	18	10.7	724	6	CB592178	EX850657
29	18	10.7	732	9	EX130112	AGENCOURT
30	18	10.7	748	7	CK352152	Danio rfr
31	18	10.7	752	9	CC923223	CK352152
32	18	10.7	755	5	BP678876	AGENCOURT
33	18	10.7	756	6	CB206666	CC923223
34	18	10.7	781	4	BG400860	BP678876
35	18	10.7	795	5	BU910572	AGENCOURT
36	18	10.7	806	8	BH360033	BG400860
37	18	10.7	806	8	BH360034	602464085
38	18	10.7	819	9	CG791244	AGENCOURT
39	18	10.7	821	9	CG490159	CH230-18J
40	18	10.7	822	9	CG043458	BH360033
41	18	10.7	857	5	BQ736159	CH230-18J
42	18	10.7	869	6	CB941331	ZMBBb028
43	18	10.7	948	6	CD255176	CG490159
44	18	10.7	1055	4	BM802977	CG791244
45	18	10.7	1084	6	CA974306	ZMBBb028

ALIGNMENTS

RESULT 1
BE950396/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
COMMENT

435 bp mRNA linear EST 29-APR-2002
UI-M-CEO-ayw-a-10-0-UI.s1 NIH BMAP_Ret3 Mus musculus cDNA clone
UI-M-CEO-ayw-a-10-0-UI 3', mRNA sequence.
BE950396
BE950396.1 GI:10589062
EST.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 435)
Bonaldo, M.F., Lennon, G. and Soares, M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)
9704477
889548
Contact: Chin, H
National Institute of Mental Health
6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD
20892-9643, USA
Tel: 301 443 1706
Fax: 301 443 9890
Email: MEST@mail.nih.gov

The sequence contained an oligo-dt track that was present in the
oligonucleotide that was used to prime the synthesis of first
strand cDNA and therefore this may represent a bonafide poly A
tail. The sequence tag present in the cDNA between the NotI site
and the oligo-dt track served to verify it as a clone from the
retina tissue cDNA library preparation: M.B. Soares Lab Clone
distribution: Researchers may obtain BMAP cDNA clones from RESEARCH
GENETICS. It should be noted that Bento Soares is generating a
small number of additional specialized non-redundant arrays of BMAP
cDNAs whose availability will be considered under appropriate and
limited collaborative arrangements. The tissue for this library was
contributed by Dr. Xin-ran Fu, Yale University School of Medicine
The following repetitive elements were found in this cDNA sequence:
1-21, >AT rich#Low complexity 169-428, >Lx9#LINE/L1
Seq primer: M13 Forward
POLYA=Yes

FEATURES
source

Location/Qualifiers
1..435
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"

```

/db_xref="taxon:10090"
/clone="UI-M-CEO-ayw-a-10-0-UI"
/dev_stage="6 weeks"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NHH_BMAP_Ret3"
/vector="pT7T3D-Pac (Pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: Eco RI; The
NHH_BMAP_Ret3 library is derived from mouse retina tissue.
For a detailed description of the library from which this
clone was derived, please visit our web site at
braines.eng.uiowa.edu. The tissue for this library was
contributed by Dr. Xin-Yuan Fu, Yale University School of
Medicine
TAG_TISSUE=adult-retina
TAG_LIB=NHH_BMAP_Ret3
TAG_SEQ=GTGAGCGCGCAC

ORIGIN
Query Match      13.1%; Score 22; DB 2; Length 435;
Best Local Similarity 100.0%; Pred.No. 0.35;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 72 CATTAAAGATCTGTCAGAAACAC 93
   |||||
Db 127 CATTAAAGATCTGTCAGAAACAC 106

RESULT 2
CK817241
LOCUS      CK817241          379 bp      mRNA      linear      EST 01-MAR-2004
DEFINITION CMA-bov OB01-009-f11 Day 0 Oviduct post-LH surge bov Bos taurus
           cDNA 3'- mRNA sequence.
ACCESSION  CK817241
VERSION    CK817241.1 GI:44834166
KEYWORDS   EST.
SOURCE     Bos taurus (cow)
ORGANISM   Bos taurus
            Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
            Bovinae; Bos.
REFERENCE  1 (bases 1 to 379)
AUTHORS   Prather,R.S., Antoniou,E., Garverick,H.A., Green,J.A., Lucy,M.C.,
            Roberts,R.M., Smith,M.F. and Youngquist,R.S.
TITLE     USDA Grant NRI-2002-03476: Bovine ESTs: Focus on Female
          Reproduction (2002)
JOURNAL   Unpublished (2002)
COMMENT   Contact: DNA Core Facility (Bovine Project)
          Animal Science - RS Prather
          University of Missouri-Columbia
          M616 Medical Sciences Bldg., Columbia, MO 65212, USA
          Tel: 573/882-0428
          Fax: 573/884-5552
          Email: bovine@net.missouri.edu
          POLYA=yes.

FEATURES
             Location/Qualifiers
             1..379
             /organism="Bos taurus"
             /mol_type="mRNA"
             /db_xref="taxon:9913"
             /clone_lib="bov"
             /note="Funding: The production of ESTs submitted in this
             project was funded by USDA Grant MRI-2002-03476 entitled
             'Bovine ESTs: Focus on Female Reproduction' to RS Prather,
             E Antoniou, HA Garverick, JA Green, MC Lucy, RM Roberts,
             MF Smith and RS Youngquist. Genetic Source: Heifers for
             the project were purchased from Circle A Ranch, Iberia, MO
             (http://www.circlea ranch.com/home.html). These heifers,
             while not registered have known Angus pedigrees going back
             at least 4 generations. Samples collected: The samples
             consisted of the following: Germinal vesicle-stage
             oocytes; in vitro derived embryos (2-cell, morula,
             blastocyst and nuclear transfer blastocyst); in vivo
             blastocysts and conceptuses (days 8, 14, 16 and 18);

```

corpora lutea (days 3, 5, 8, 14, 16, 18 and 35); ovarian follicles (days 0, non-recruited, recruited, early selected and preovulatory); oviduct (days 0, 3 and 5); endometrium (days 5, 8, 14, 16, 18 and 35); and placenta/embryo from day 35 conceptuses. Expanded descriptions of how the tissues were collected can be found at the following URL: <http://genome.rnet.missouri.edu/Bovine/Methods.html>. Library construction (Standard Protocol): All procedures have been described in detail elsewhere (Soares et al., 1994; Bonaldo et al., 1996; Jiang et al., 2001). Total cellular RNA from each sample was isolated by using STAT-60 reagent (Tel-Test, Friendswood, TX) and the poly(A)⁺ RNA was obtained by two rounds of purification with the Oligotex mRNA isolation kit (Qiagen) according to the manufacturer's instructions. The libraries were described by the manufacturer's instructions provided with the SuperScript Plasmid System (Invitrogen, cat. no. 18248-013). Briefly, lmcg of poly(A)⁺ RNA was annealed at 37 degrees with lmcg of NotI-tag-dT18 oligonucleotide (GCTGTCGCGCGCG-tag-T18) and reverse transcribed at c37 degrees with Superscript II (Invitrogen) reverse transcriptase (Jiang et al., 2001). The 'tag' represents a tissue/stage-specific ten-base sequence identifier (<http://genome.uiowa.edu/pubsoft/software.html>) present in the oligonucleotide used to prime first-strand synthesis. Second strand synthesis was performed with T4 DNA Polymerase in the presence of DNA ligase and RNase H. After second strand synthesis, the double-stranded cDNAs were ligated to SalI adapters (Invitrogen-Life Technologies) and digested with NotI. The cDNAs were size selected by passage through cDNA size fractionation columns (Invitrogen-Life Technologies). The cDNAs derived from each developmental stage of a particular tissue were mixed on an equimolar basis and ligated directionally into the NotI and SalI sites of the pCMV-SPORT6 vector (Invitrogen). After ligation of the inserts, the plasmids were electroporated into DH10B bacteria. Library that was constructed (PCR Protocol): The amount of mRNA that was recovered from oocytes and embryos was quite limiting and was not sufficient for library production with the standard protocol. Therefore, PCR-based protocol was utilized for producing libraries from sources in which the amount of extracted mRNA was small (oocytes and embryos). Poly-A RNA was isolated by using the MicroPoly(A) Pure kit from Ambion (cat. # 1918). The mRNA was reverse transcribed with a NotI-tag-dT18 oligonucleotide and a SMART oligonucleotide (Clontech) modified to contain a SalI site to generate full-length cDNA with a sequence complementary to the SMART oligonucleotide. Sequences within the SMART and dt oligonucleotides were used as primers to amplify the cDNAs by PCR with pfu turbo purified, digested with NotI and SalI and size fractionated by using Chroma Spin-1000 columns (Clontech). Purified cDNA from each PCR reaction was quantitated and mixed on an equimolar basis for ligation into the pCMV-SPORT6 vector. Preliminary Library Characterization: Randomly chosen clones from each library were analyzed by restriction digestion to determine average insert size (96 clones) and by sequencing (~4 96-well plates) to confirm library quality (e.g. the presence of short polyA tails, genomic DNA contamination (must be <1%), ribosomal RNA clones (must be <1%), etc.) and to provide a sequence database representing the predominant clones in each library. The clones were sequenced at the University of Missouri-Columbia DNA Core Facility. After production of the libraries, equal numbers of recombinants from each library were pooled to produce a single mixed library (mega-library) for more extensive sequencing. Bioinformatics work was performed by GK Springer's bioinformatics group (Spollen WG, Topinka CM, Khambati AA)

in Computer Science at the University of Missouri-Columbia. Clone Requests: Requests for clones should be made to the Director of the University of Missouri DNA Core facility at: bovine@net.missouri.edu. Bonaldo MF, Lennon G, Soares MB. Normalization and Subtraction: Two approaches to facilitate gene discovery. Genome Res, 1996; 6:791-806. Jiang H, Bivens NJ, Ries JS, Whitworth KM, Green JA, Forrester LJ, Springer GK, Didion BA, Mathialagan N, Prather RS, Lucy MC (2001) Constructing cDNA libraries with fewer clones that contain long poly(dA) tails. Biotechniques 31:38-42. Soares MB, MF Banaldo, P Jelene, L Su, L Lawton, A Efstratiadis. 1994. Construction and characterization of a normalized cDNA library. Proc Natl Acad Sci, 91:9228-9232. TAG_TISSUE=Day 0 Oviduct post-LH surge TAG_SEQ=GTGACCGATG"

ORIGIN

Query Match 12.5%; Score 21; DB 7; Length 379;
Best Local Similarity 100.0%; Pred. No. 1.3;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 72 CATTAAAGATCTGTCAGAAACA 92
|||||
DB 343 CATTAAAGATCTGTCAGAAACA 363

RESULT 3

AJ814218/c AJ814218 665 bp mRNA linear EST 13-SEP-2004
LOCUS
DEFINITION
ACCESSION
VERSION
AJ814218.1 GI:51881694
KEYWORDS
EST.
SOURCE
Bos sp.
ORGANISM

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
McGuire, K. and Glass, E.J.
Unpublished ESTs, McGuire and Glass
Contact: McGuire K
Genomics and Genetics
Roslin Institute
Roslin, Midlothian, EH25 9PS, UNITED KINGDOM
Single pass sequencing. Bases called and trimmed with phred v0.020425.c. Vector identified by cross_match with the -minscore 20 and -minmatch 12 options. Vector:pBluescriptII(SK+) R. Site 1:EcoRV(loest) R. Site 2:NotI Seq Primer: T7 Normalised library constructed from pooled monocytes from Bos taurus (Holstein) and Bos indicus (Sahiwal) cattle subjected to various stimuli, including infection with the protozoan parasite Theileria annulata.

1 (bases 1 to 665)

/organism="Bos sp."
/mol_type="mRNA"
/db_xref="taxon:29061"
/clone="C0005198j2"
/tissue_type="blood"
/cell_type="bovine monocyte"
/clone_lib="KN206"

/note="Vector: pBluescriptII(SK+); Site 1: EcoRV(loest); Site 2: NotI; Normalised library constructed from pooled monocytes from Bos taurus (Holstein) and Bos indicus (Sahiwal) cattle subjected to various stimuli, including infection with the protozoan parasite Theileria annulata"

FEATURES

source
1..665
/organism="Bos sp."
/mol_type="mRNA"
/db_xref="taxon:29061"
/clone="C0005198j2"
/tissue_type="blood"
/cell_type="bovine monocyte"
/clone_lib="KN206"

/note="Vector: pBluescriptII(SK+); Site 1: EcoRV(loest); Site 2: NotI; Normalised library constructed from pooled monocytes from Bos taurus (Holstein) and Bos indicus (Sahiwal) cattle subjected to various stimuli, including infection with the protozoan parasite Theileria annulata"

ORIGIN

Query Match 12.5%; Score 21; DB 1; Length 665;
Best Local Similarity 100.0%; Pred. No. 1.3;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 72 CATTAAAGATCTGTCAGAAACA 92
|||||
DB 360 CATTAAAGATCTGTCAGAAACA 340

RESULT 4

CB166163
LOCUS
DEFINITION
CB166163 674 bp mRNA linear EST 30-JAN-2003
X186030104860.R1 CSEQFXN32 hypothalamus Bos taurus cDNA, mRNA
sequence.
CB166163
ACCESSION
VERSION
CB166163.1 GI:28152288
KEYWORDS
EST.
SOURCE
Bos taurus (cow)
ORGANISM

Bos taurus

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
1 (bases 1 to 674)
Adelson, D.L. and Gill, C.A.
Bovine ESTs (Adelson and Gill)
Unpublished (2003)
Contact: David L. Adelson
Animal Breeding and Genetics
Texas A&M University
Animal Science Dept., TAMU-2471, College Station, TX 77843-2471,
USA
Tel: 9798452616
Fax: 9798456970
Email: david.adelson@tamu.edu.

Location/Qualifiers
1..674
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
/tissue_type="hypothalamus"
/clone_lib="CSEQFXN32 hypothalamus"
/note="Organ: hypothalamus; Vector: pBluescript SK+;
Site 1: NotI; Site 2: EcoRI; sequence 5' of the insert
(5'-NNN...NNNinsert)
GCGAATTGGAGCTCCACCGCGTGGCGCGCGCTCGAG. Sequence 3' of
the inserts (AAGAATTCGATCAAGCTATCGATACCGTCCGCTCGAG.
normalized Rd 1 library, sequenced 3' with M13R primer."

1 (bases 1 to 728)

/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
/tissue_type="hypothalamus"
/clone_lib="CSEQFXN32 hypothalamus"
/note="Organ: hypothalamus; Vector: pBluescript SK+;
Site 1: NotI; Site 2: EcoRI; sequence 5' of the insert
(5'-NNN...NNNinsert)
GCGAATTGGAGCTCCACCGCGTGGCGCGCGCTCGAG. Sequence 3' of
the inserts (AAGAATTCGATCAAGCTATCGATACCGTCCGCTCGAG.
normalized Rd 1 library, sequenced 3' with M13R primer."

1 (bases 1 to 728)

Sonstegard, T.S., Van Tassel, C.P., Matukumalli, L.K., Harhay,
G.P., Bosak, S., Rubenfield, M. and Gaabbarre, L.C.
Production of EST from cDNA libraries derived from immunologically
activated bovine gut

1 (bases 1 to 728)

Sonstegard, T.S., Van Tassel, C.P., Matukumalli, L.K., Harhay,
G.P., Bosak, S., Rubenfield, M. and Gaabbarre, L.C.
Production of EST from cDNA libraries derived from immunologically
activated bovine gut

1 (bases 1 to 728)

Sonstegard, T.S., Van Tassel, C.P., Matukumalli, L.K., Harhay,
G.P., Bosak, S., Rubenfield, M. and Gaabbarre, L.C.
Production of EST from cDNA libraries derived from immunologically
activated bovine gut

1 (bases 1 to 728)

Sonstegard, T.S., Van Tassel, C.P., Matukumalli, L.K., Harhay,
G.P., Bosak, S., Rubenfield, M. and Gaabbarre, L.C.
Production of EST from cDNA libraries derived from immunologically
activated bovine gut

1 (bases 1 to 728)

Sonstegard, T.S., Van Tassel, C.P., Matukumalli, L.K., Harhay,
G.P., Bosak, S., Rubenfield, M. and Gaabbarre, L.C.
Production of EST from cDNA libraries derived from immunologically
activated bovine gut

1 (bases 1 to 728)

Sonstegard, T.S., Van Tassel, C.P., Matukumalli, L.K., Harhay,
G.P., Bosak, S., Rubenfield, M. and Gaabbarre, L.C.
Production of EST from cDNA libraries derived from immunologically
activated bovine gut

1 (bases 1 to 728)

Sonstegard, T.S., Van Tassel, C.P., Matukumalli, L.K., Harhay,
G.P., Bosak, S., Rubenfield, M. and Gaabbarre, L.C.
Production of EST from cDNA libraries derived from immunologically
activated bovine gut

1 (bases 1 to 728)

Sonstegard, T.S., Van Tassel, C.P., Matukumalli, L.K., Harhay,
G.P., Bosak, S., Rubenfield, M. and Gaabbarre, L.C.
Production of EST from cDNA libraries derived from immunologically
activated bovine gut

1 (bases 1 to 728)

Sonstegard, T.S., Van Tassel, C.P., Matukumalli, L.K., Harhay,
G.P., Bosak, S., Rubenfield, M. and Gaabbarre, L.C.
Production of EST from cDNA libraries derived from immunologically
activated bovine gut

1 (bases 1 to 728)

Sonstegard, T.S., Van Tassel, C.P., Matukumalli, L.K., Harhay,
G.P., Bosak, S., Rubenfield, M. and Gaabbarre, L.C.
Production of EST from cDNA libraries derived from immunologically
activated bovine gut

1 (bases 1 to 728)

Sonstegard, T.S., Van Tassel, C.P., Matukumalli, L.K., Harhay,
G.P., Bosak, S., Rubenfield, M. and Gaabbarre, L.C.
Production of EST from cDNA libraries derived from immunologically
activated bovine gut

1 (bases 1 to 728)

Sonstegard, T.S., Van Tassel, C.P., Matukumalli, L.K., Harhay,
G.P., Bosak, S., Rubenfield, M. and Gaabbarre, L.C.
Production of EST from cDNA libraries derived from immunologically
activated bovine gut

1 (bases 1 to 728)

Sonstegard, T.S., Van Tassel, C.P., Matukumalli, L.K., Harhay,
G.P., Bosak, S., Rubenfield, M. and Gaabbarre, L.C.
Production of EST from cDNA libraries derived from immunologically
activated bovine gut

1 (bases 1 to 728)

Sonstegard, T.S., Van Tassel, C.P., Matukumalli, L.K., Harhay,
G.P., Bosak, S., Rubenfield, M. and Gaabbarre, L.C.
Production of EST from cDNA libraries derived from immunologically
activated bovine gut

1 (bases 1 to 728)

Sonstegard, T.S., Van Tassel, C.P., Matukumalli, L.K., Harhay,
G.P., Bosak, S., Rubenfield, M. and Gaabbarre, L.C.
Production of EST from cDNA libraries derived from immunologically
activated bovine gut

1 (bases 1 to 728)

Sonstegard, T.S., Van Tassel, C.P., Matukumalli, L.K., Harhay,
G.P., Bosak, S., Rubenfield, M. and Gaabbarre, L.C.
Production of EST from cDNA libraries derived from immunologically
activated bovine gut

1 (bases 1 to 728)

Sonstegard, T.S., Van Tassel, C.P., Matukumalli, L.K., Harhay,
G.P., Bosak, S., Rubenfield, M. and Gaabbarre, L.C.
Production of EST from cDNA libraries derived from immunologically
activated bovine gut

1 (bases 1 to 728)

Sonstegard, T.S., Van Tassel, C.P., Matukumalli, L.K., Harhay,
G.P., Bosak, S., Rubenfield, M. and Gaabbarre, L.C.
Production of EST from cDNA libraries derived from immunologically
activated bovine gut

1 (bases 1 to 728)

Sonstegard, T.S., Van Tassel, C.P., Matukumalli, L.K., Harhay,
G.P., Bosak, S., Rubenfield, M. and Gaabbarre, L.C.
Production of EST from cDNA libraries derived from immunologically
activated bovine gut

1 (bases 1 to 728)

Sonstegard, T.S., Van Tassel, C.P., Matukumalli, L.K., Harhay,
G.P., Bosak, S., Rubenfield, M. and Gaabbarre, L.C.
Production of EST from cDNA libraries derived from immunologically
activated bovine gut

1 (bases 1 to 728)

Sonstegard, T.S., Van Tassel, C.P., Matukumalli, L.K., Harhay,
G.P., Bosak, S., Rubenfield, M. and Gaabbarre, L.C.
Production of EST from cDNA libraries derived from immunologically
activated bovine gut

1 (bases 1 to 728)

Sonstegard, T.S., Van Tassel, C.P., Matukumalli, L.K., Harhay,
G.P., Bosak, S., Rubenfield, M. and Gaabbarre, L.C.
Production of EST from cDNA libraries derived from immunologically
activated bovine gut

1 (bases 1 to 728)

Sonstegard, T.S., Van Tassel, C.P., Matukumalli, L.K., Harhay,
G.P., Bosak, S., Rubenfield, M. and Gaabbarre, L.C.
Production of EST from cDNA libraries derived from immunologically
activated bovine gut

1 (bases 1 to 728)

Sonstegard, T.S., Van Tassel, C.P., Matukumalli, L.K., Harhay,
G.P., Bosak, S., Rubenfield, M. and Gaabbarre, L.C.
Production of EST from cDNA libraries derived from immunologically
activated bovine gut

1 (bases 1 to 728)

Sonstegard, T.S., Van Tassel, C.P., Matukumalli, L.K., Harhay,
G.P., Bosak, S., Rubenfield, M. and Gaabbarre, L.C.
Production of EST from cDNA libraries derived from immunologically
activated bovine gut

1 (bases 1 to 728)

Sonstegard, T.S., Van Tassel, C.P., Matukumalli, L.K., Harhay,
G.P., Bosak, S., Rubenfield, M. and Gaabbarre, L.C.
Production of EST from cDNA libraries derived from immunologically
activated bovine gut

1 (bases 1 to 728)

Sonstegard, T.S., Van Tassel, C.P., Matukumalli, L.K., Harhay,
G.P., Bosak, S., Rubenfield, M. and Gaabbarre, L.C.
Production of EST from cDNA libraries derived from immunologically
activated bovine gut

1 (bases 1 to 728)

Sonstegard, T.S., Van Tassel, C.P., Matukumalli, L.K., Harhay,
G.P., Bosak, S., Rubenfield, M. and Gaabbarre, L.C.
Production of EST from cDNA libraries derived from immunologically
activated bovine gut

1 (bases 1 to 728)

Sonstegard, T.S., Van Tassel, C.P., Matukumalli, L.K., Harhay,
G.P., Bosak, S., Rubenfield, M. and Gaabbarre, L.C.
Production of EST from cDNA libraries derived from immunologically
activated bovine gut

1 (bases 1 to 728)

Sonstegard, T.S., Van Tassel, C.P., Matukumalli, L.K., Harhay,
G.P., Bosak, S., Rubenfield, M. and Gaabbarre, L.C.
Production of EST from cDNA libraries derived from immunologically
activated bovine gut

1 (bases 1 to 728)

Sonstegard, T.S., Van Tassel, C.P., Matukumalli, L.K., Harhay,
G.P., Bosak, S., Rubenfield, M. and Gaabbarre, L.C.
Production of EST from cDNA libraries derived from immunologically
activated bovine gut

```

JOURNAL
COMMENT
Unpublished (2004)
Contact: Tad S. Sonstegard
Bovine Functional Genomics Laboratory
Animal and Natural Resources Institute
Bldg. 200 Rm2A BARC-East, Beltsville, MD 20705, USA
Tel: 3015048416
Fax: 3015048414
Email: tads@nri.barc.usda.gov
Single pass sequencing. Bases called and trimmed with phred
0.000925 using options -trim_alt "-trim.fasta. Vector identified
by cross_match using options -minmatch 12 -minscore 12
Plate: 26 row: I column: 08
Seq primer: CCCAGTCACGAGTGTGTAAGCG
High quality sequence stop: 728.

FEATURES
source
Location/Qualifiers
1..728
/organism="Bos taurus"
/mol_type="mRNA"
/strains="Holstein"
/db_xref="taxon:9913"
/clone="108OV26_108"
/sex="Male"
/tissue_type="Pooled"
/dev_stage="Multiple"
/lab_host="DH10B T1 phage resistant"
/clone_lib="BARC 108OV"
/notes="Organ: Small Intestine; Vector: pBgen-1; Site:1:
EcoRI; Site 2: NotI; Equimolar amounts of mRNA extracted
from proximal jejunums of 18 and 21 wk old steers, and
distal ileums of 14 day old calves. proximal jejunum
exposed to C. oncophora for 3 and 6 weeks, and distal
ileum exposed to C. parvum for 7 days"

ORIGIN
Query Match 12.5%; Score 21; DB 7; Length 728;
Best Local Similarity 100.0%; Pred. No. 1.3;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 72 CATTAGATCTCTCAGAAACA 92
Db 499 CATTAGATCTCTCAGAAACA 479

RESULT 6
AW361823
LOCUS
DEFINITION
PM0-CT0263-021299-006-b02 CT0263 Homo sapiens cDNA, mRNA sequence.
ACCESSION
AW361823
VERSION
AW361823.1 GI:6866473
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 580)
HCGP http://www.ludwig.org.br/ORESTES
The FAPESP/LICR Human Cancer Genome Project
Unpublished (1999)
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?ti=PM0&t2=PM0-CT0263-
021299-006-b02&t3=1999-12-02&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 3
High quality sequence stop: 148.

FEATURES
source
Location/Qualifiers
1..580
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="CT0263"
/notes="Organ: colon; Vector: puc18; Site_1: SmaI; Site_2:
SmaI; A mini-library was made by cloning products derived
from ORESTES PCR (U.S. Letters Patent application No.
196,716 - Ludwig Institute for Cancer Research) profiles
into the pUC 18 vector. Reverse transcription of tissue
mRNA and cDNA amplification were performed under low
stringency conditions."

ORIGIN
Query Match 11.3%; Score 19; DB 2; Length 580;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 120 CTTCATGCTGCTGTGCAA 138
Db 98 CTTCATGCTGCTGTGCAA 116

RESULT 7
B10170
LOCUS
DEFINITION
F14D20-SP6 IGF Arabidopsis thaliana genomic clone F14D20, genomic
survey sequence.
ACCESSION
B10170
VERSION
B10170.1 GI:2091289
KEYWORDS
GSS.
SOURCE
Arabidopsis thaliana (thale cress)
ORGANISM
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 1101)
Feng,J., Dewar,K., Buehler,E., Kim,C., Li,Y., Shinn,P., Sun,H. and
Ecker,J.
BAC End Sequences at ATGC
Unpublished (1997)
Other GSSs: F14D20-T7
Contact: Ecker J.
Arabidopsis Thaliana Genome Center
University of Pennsylvania
Dept. of Biology, University of Pennsylvania, Philadelphia, PA
19104
Tel: 215-898-9384
Fax: 215-898-8780
Email: jecker@atgenome.bio.upenn.edu
Seq primer: Sp6
Class: BAC ends
High quality sequence start: 121
High quality sequence stop: 802.

FEATURES
source
Location/Qualifiers
1..1101
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/ecotype="Columbia"
/db_xref="taxon:3702"
/clone="F14D20"
/sex="hermaphrodite"
/clone_lib="IGF"
/notes="Vector: BelOBACII; Site_1: EcoRI; Site_2: EcoRI;
Produced by Thomas Altmann"

ORIGIN
Query Match 11.3%; Score 19; DB 8; Length 1101;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 20.0 and subtraction to Rot = 370.0. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGATTCGAGTTAATAATTAATTCCTCCCCCCCC 3']. cDNA was cleaved with XhoI and BamHI. Vector: a modified plasmid KS(+) after bulk excision from Lambda FLC I."

QY 137 AATTTCTGAAGGATGAGAA 155
 |||||
Db 405 AATTTCTGAAGGATGAGAA 423

RESULT 8
BB586775 263 bp mRNA linear EST 30-NOV-2000
LOCUS musculus cDNA clone 9570005A14 5', mRNA sequence.
DEFINITION BB586775
ACCESSION BB586775.1 GI:11483319
VERSION EST.
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Mus musculus
REFERENCE Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 263)
AUTHORS Aizawa, K., Akahira, S., Akimura, T., Arai, A., Arakawa, T., Carninci, P., Hanagaki, T., Hayatsu, N., Hiraoka, T., Hirozane, T., Hodojima, Y., Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kawai, J., Kojima, Y., Konno, H., Kuakabe, M., Matsuyama, T., Miyazaki, A., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Okazaki, Y., Okido, T., Owa, C., Sakai, C., Sakai, K., Sasaki, D., Sato, K., Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagawa, A., Takahashi, F., Tanaka, T., Toya, T., Watahiki, A., Yamamura, T., Yasunishi, A., Yoshida, K., Yoshiki, A., Muramatsu, M. and Hayashizaki, Y.
TITLE RIKEN Mouse ESTs (Aizawa, K. et al. 2000)
JOURNAL Unpublished (2000)
COMMENT Contact: Yoshihide Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 Tel: 81-45-503-9222
 Fax: 81-45-503-9216
 Email: genome-resgsc.riken.jp, URL: http://genome.gsc.riken.jp/
 Carninci, P., Nishiyama, Y., Westover, A., Itoh, M., Nagaoka, S., Sasaki, N., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 Thermostabilization and thermoactivation of thermolabile enzymes by trehalose and its application for the synthesis of full length cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)
 Itoh, M., Kitesunai, T., Akiyama, J., Shibata, K., Izawa, M., Kawai, J., Tomaru, Y., Carninci, P., Shibata, Y., Ozawa, Y., Muramatsu, M., Okazaki, Y. and Hayashizaki, Y.
 Automated filtration-based high-throughput plasmid preparation system. Genome Res. 9 (5), 463-470 (1999)
 Carninci, P. and Hayashizaki, Y.
 High-efficiency full-length cDNA cloning. Methods Enzymol. 303, 19-44 (1999)
 Please visit our web site (http://genome.rtc.riken.go.jp) for further details.

FEATURES
 source
 Location/Qualifiers
 1..263
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /clone="9570005A14"
 /sex="male"
 /tissue_type="urinary bladder"
 /dev_stage="adult"
 /lab_host="DH10B"
 /clone_lib="RIKEN full-length enriched, adult male urinary bladder"
 /note="Site 1: SalI; Site 2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5' GAGAGAGAGGATCCAGAGCTCTTTTCTTTTCTTTTCTTTT 3'], cDNA was

ORIGIN
 Query Match 10.7%; Score 18; DB 2; Length 263;
 Best Local Similarity 100.0%; Pred. No. 67;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 133 TGTGAATTTCTGAAGGAT 150
 |||||
Db 168 TGTGAATTTCTGAAGGAT 185

RESULT 9
BG603087/c 313 bp mRNA linear EST 14-AUG-2001
LOCUS EST502177 Plasmodium yoelii sporozoite cDNA Plasmodium yoelii cDNA
DEFINITION BG603087
ACCESSION BG603087.1 GI:15153101
VERSION EST.
KEYWORDS Plasmodium yoelii
SOURCE Plasmodium yoelii
ORGANISM Eukaryota; Alveolata; Apicomplexa; Haemosporidia; Plasmodium.
REFERENCE 1 (bases 1 to 313)
AUTHORS Kappe, S.H.I., Gardner, M.J., Brown, S.M., Ross, J., Matuschewski, K., Ribeiro, J.M., Adams, J.H., Quackenbush, J., Cho, J., Carucci, D.J., Hoffman, S.B. and Nussenzweig, V.
TITLE Exploring the transcriptome of the malaria sporozoite stage
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 98 (17), 9895-9900 (2001)
MEDLINE 11493655
PUBMED 11493655
COMMENT Contact: Malcolm J. Gardner
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Drive, Rockville, MD 20850, USA
 Tel: 301 838 3519
 Fax: 301 838 0208
 Email: gardner@tigr.org
 Request for clones, please contact: Stefan Kappe, Division, kappes01@popmail.med.nyu.edu Michael Heidelberger, Division, Department of Pathology New York University School of Medicine.
 Location/Qualifiers
 1..313
 /organism="Plasmodium yoelii"
 /mol_type="mRNA"
 /strain="17XNL"
 /db_xref="taxon:5961"
 /clone="PYCDN74"
 /dev_stage="sporozoites from salivary gland"
 /lab_host="E. coli TOP10"
 /clone_lib="Plasmodium yoelii sporozoite cDNA"
 /note="Vector: pCR4; TA cloning; Plasmodium yoelii sporozoite cDNA library from salivary gland sporozoites 14 days post-infection"

ORIGIN
 Query Match 10.7%; Score 18; DB 4; Length 313;
 Best Local Similarity 100.0%; Pred. No. 67;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 125 AATGCTGGTGTGAATTC 142
 |||||
Db 218 AATGCTGGTGTGAATTC 201

```

RESULT 10
BZ686719/c
LOCUS      BZ686719          360 bp      DNA          linear      GSS 05-FEB-2003
DEFINITION PUBU54TD.ZM.0.6_1.0_KB_Zea_mays_genomic_clone_ZMUBTA019J11,
            genomic survey sequence.
ACCESSION  BZ686719
VERSION    BZ686719.1  GI:28247227
KEYWORDS   GSS.
SOURCE     Zea mays
ORGANISM   Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE  1 (bases 1 to 360)
AUTHORS   Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
            Renick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
            Bennetzen,J.
TITLE     Maize Genomic Consortium
JOURNAL   Unpublished (2003)
COMMENT   Contact: Cathy Whitelaw
            TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA
            Tel: 301-838-5843
            Fax: 301-838-0208
            Email: whitelaw@tigr.org
            Seq primer: 1F
            Class: sheared ends.
            Location/Qualifiers
                1..360
                /organism="Zea mays"
                /mol_type="genomic DNA"
                /strain="B73"
                /db_xref="taxon:4577"
                /clone="ZMUBTA019J11"
                /clone_lib="ZM 0.6_1.0_KB"
                /note="Vector: PCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high
                Cot selected genomic DNA library"

ORIGIN
Query Match      10.7%; Score 18; DB 8; Length 360;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 75 TAAGATCTGTCAGAAACA 92
Db 57 TAAGATCTGTCAGAAACA 40

RESULT 11
AQ405655/c
LOCUS      AQ405655          423 bp      DNA          linear      GSS 13-MAR-1999
DEFINITION HS_5049_B1.G07_SP6E_RPCI-11_Human_Male_BAC_Library_Homo_sapiens
            genomic clone Plates=625 Col=13 Row=N, genomic survey sequence.
ACCESSION  AQ405655
VERSION    AQ405655.1  GI:4415643
KEYWORDS   GSS.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE  1 (bases 1 to 423)
AUTHORS   Mahaitas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
            Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
            Hood,L.
TITLE     Sequence-tagged connectors: A sequence approach to mapping and
            scanning the human genome
JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
MEDLINE   99380589
PUBMED    10449764
COMMENT   Contact: Mahairas GG, Wallace JC, Hood L
            High Throughput Sequencing Center
            University of Washington
            401 Queen Anne Avenue North, Seattle, WA 98109, USA

```

```

Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Clones are derived from the human BAC library RPCI-11. For BAC
library availability, please contact Pieter de Jong
(pieter@dejong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm)
or from Research Genetics (Info@resgen.com). BAC end Web Server:
http://www.htsc.washington.edu
Plate: 625 row: N column: 13
Seq primer: SP6
Class: BAC ends
High quality sequence stop: 423.
            Location/Qualifiers
                1..423
                /organism="Homo sapiens"
                /mol_type="genomic DNA"
                /db_xref="taxon:9606"
                /clone="Plate=625 Col=13 Row=N"
                /sex="male"
                /clone_lib="RPCI-11 Human Male BAC Library"
                /note="Vector: pBACe3.6; Site 1: EcoRI; Site 2: EcoRI;
                Male blood DNA was isolated from one randomly chosen donor
                and partially digested with a combination of EcoRI and
                EcoRI Methylase. Size selected DNA was cloned into the
                pBACe3.6 vector at EcoRI sites"

ORIGIN
Query Match      10.7%; Score 18; DB 8; Length 423;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 132 GTGTGAATTTCTGAAGGA 149
Db 260 GTGTGAATTTCTGAAGGA 243

RESULT 12
CF794539
LOCUS      CF794539          466 bp      mRNA          linear      EST 21-OCT-2003
DEFINITION 890065_MARC_4PIG_Sus_scrofa_cdna_5', mRNA sequence.
ACCESSION  CF794539
VERSION    CF794539.1  GI:37799112
KEYWORDS   EST.
SOURCE     Sus scrofa (pig)
ORGANISM   Sus scrofa
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE  1 (bases 1 to 466)
AUTHORS   Smith,T.P.L., Preking,B.A., Ford,J.J., Vallet,J.L., Wise,T.A.,
            Nonneman,D.J., Wray,J.E. and Keele,J.W.
TITLE     Porcine EST collection using a normalized library constructed from
            embryos representing early developmental stages
JOURNAL   Unpublished (2003)
COMMENT   Contact: Smith TPL
            USDA, ARS, US Meat Animal Research Center
            PO Box 166, Clay Center, NE 68933-0166, USA
            Tel: 402 762 4366
            Fax: 402 762 4390
            Email: smith@email.marc.usda.gov
            Single pass sequencing. Bases called with phred v0.020425.c and
            trimmed with the aid of the trim_alt option. Vector identified with
            cross_match v0.990329.
            Plate: TW8014 row: E column: 12
            Seq primer: GTAATACGACTCACTATAGG.
            Location/Qualifiers
                1..466
                /organism="Sus scrofa"
                /mol_type="mRNA"
                /db_xref="taxon:9823"
                /tissue_type="pooled"
                /lab_host="DH10B"
                /clone_lib="MARC 4PIG"

```


/note="Vector: pCDNA3.1; Site 1: EcoRI; Site 2: NotI; library made with combined RNA from day-10, day-13, day-15, day-25, and day-30 whole embryos."

ORIGIN

Query Match 10.7%; Score 18; DB 7; Length 466;
Best Local Similarity 100.0%; Pred. No. 69;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 136 GAATTTCTGAAGGATGAG 153

Db 108 GAATTTCTGAAGGATGAG 125

RESULT 13

AZ644277

LOCUS

DEFINITION IM0508H1F Mouse 10kb plasmid UUGC1M library Mus musculus genomic clone UUGC1M0508H1 F, genomic survey sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

CONTACT: Robert B. Weiss

UNIVERSITY OF UTAH

RM. 308, BIOMEDICAL POLYMERS RESEARCH BLDG., 20 S. 2030 E., SLC, UT

84112, USA

TEL: 801 585 5606

FAX: 801 585 7177

EMAIL: ddunn@genetics.utah.edu

INSERT LENGTH: 10000

PLATE: 0508

ROW: H

COLUMN: 11

SEQ PRIMER: CGTTGTAAACGACGCCAGT

CLASS: plasmid ends

HIGH QUALITY SEQUENCE STOP: 466.

LOCATION/QUALIFIERS

1..466

/organism="Mus musculus"

/mol_type="genomic DNA"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="UUGC1M0508H1"

/sex="Male"

/lab_host="E. Coli strain XL10-Gold, Ti-resistant, F-"

/clone_lib="Mouse 10kb plasmid UUGC1M library"

/note="Vector: PWD42nv; Purified genomic DNA from M. Laboratory Mouse DNA Resource

(http://www.jax.org/resources/documents/dnares/). The DNA

was hydrodynamically sheared by repeated passage through a

0.005 inch orifice at constant velocity. The sheared DNA

was blunt end-repaired with T4 DNA polymerase and T4

polynucleotide kinase. Adaptor oligonucleotides were

ligated to the blunt ends in high molar excess. The

adapted DNA was purified and size-selected for a 9.5 to

10.5 kb range using preparative agarose gel

electrophoresis. Vector DNA was prepared from a derivative

of PWD42 (gi|4732114|gb|AF129072.1), a copy-number

inducible derivative of plasmid R1. The vector was ligated

with adaptors complementary to the insert adaptors and

purified. The sheared, adapted mouse DNA was annealed to

ORIGIN

Query Match 10.7%; Score 18; DB 8; Length 466;
Best Local Similarity 100.0%; Pred. No. 69;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 133 TGTGAATTTCTGAAGGAT 150

Db 102 TGTGAATTTCTGAAGGAT 119

RESULT 14

BG603242/c

LOCUS

DEFINITION BG603242 477 bp mRNA linear EST 14-AUG-2001 clone PYCDP45, mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

COMMENT

CONTACT: Malcolm J. Gardner

DEPARTMENT OF EUKARYOTIC GENOMICS

THE INSTITUTE FOR GENOMIC RESEARCH

9712 MEDICAL CENTER DRIVE, ROCKVILLE, MD 20850, USA

TEL: 301 838 3519

FAX: 301 838 0208

EMAIL: gardner@igr.org

REQUEST FOR CLONES, PLEASE CONTACT: Stefan Kappe, Division,

kappes01@popmail.med.nyu.edu Michael Heidelberger, Division,

Department of Pathology New York University School of Medicine.

LOCATION/QUALIFIERS

1..477

/organism="Plasmodium yoelii"

/mol_type="mRNA"

/strain="17XNL"

/db_xref="taxon:5861"

/clone="PYCDP45"

/dev_stage="sporozoites from salivary gland"

/lab_host="E. coli TOP10"

/clone_lib="Plasmodium yoelii sporozoite cDNA"

/note="Vector: pCR4; TA cloning; Plasmodium yoelii

sporozoite cDNA library from salivary gland sporozoites 14

days post-infection"

ORIGIN

Query Match 10.7%; Score 18; DB 4; Length 477;

Best Local Similarity 100.0%; Pred. No. 69;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 125 AATGCTGCTGGAATTC 142

Db 221 AATGCTGCTGGAATTC 204

RESULT 15

PT003115U/c

LOCUS

DEFINITION PT003115U 483 bp DNA linear GSS 29-MAY-2003

Paramesicium tetraurelia sequence M03E08u of the end of plasmid

PT003115, genomic survey sequence.

ACCESSION

AL446451

adapted vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

ORIGIN

Query Match 10.7%; Score 18; DB 8; Length 466;
Best Local Similarity 100.0%; Pred. No. 69;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 133 TGTGAATTTCTGAAGGAT 150

Db 102 TGTGAATTTCTGAAGGAT 119

RESULT 14

BG603242/c

LOCUS

DEFINITION BG603242 477 bp mRNA linear EST 14-AUG-2001 clone PYCDP45, mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

COMMENT

CONTACT: Malcolm J. Gardner

DEPARTMENT OF EUKARYOTIC GENOMICS

THE INSTITUTE FOR GENOMIC RESEARCH

9712 MEDICAL CENTER DRIVE, ROCKVILLE, MD 20850, USA

TEL: 301 838 3519

FAX: 301 838 0208

EMAIL: gardner@igr.org

REQUEST FOR CLONES, PLEASE CONTACT: Stefan Kappe, Division,

kappes01@popmail.med.nyu.edu Michael Heidelberger, Division,

Department of Pathology New York University School of Medicine.

LOCATION/QUALIFIERS

1..477

/organism="Plasmodium yoelii"

/mol_type="mRNA"

/strain="17XNL"

/db_xref="taxon:5861"

/clone="PYCDP45"

/dev_stage="sporozoites from salivary gland"

/lab_host="E. coli TOP10"

/clone_lib="Plasmodium yoelii sporozoite cDNA"

/note="Vector: pCR4; TA cloning; Plasmodium yoelii

sporozoite cDNA library from salivary gland sporozoites 14

days post-infection"

ORIGIN

Query Match 10.7%; Score 18; DB 4; Length 477;

Best Local Similarity 100.0%; Pred. No. 69;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 125 AATGCTGCTGGAATTC 142

Db 221 AATGCTGCTGGAATTC 204

RESULT 15

PT003115U/c

LOCUS

DEFINITION PT003115U 483 bp DNA linear GSS 29-MAY-2003

Paramesicium tetraurelia sequence M03E08u of the end of plasmid

PT003115, genomic survey sequence.

ACCESSION

AL446451

```

VERSION      AL446451.1  GI:11121726
KEYWORDS     GSS.
SOURCE       Parametium tetraurelia
ORGANISM     Parametium tetraurelia
              Eukaryota; Alveolata; Ciliophora; Oligohymenophorea; Peniculida;
              Parametium.
REFERENCE    1
AUTHORS      Keller,A.M. and Cohen,J.
TITLE        An indexed genomic library for Parametium complementation cloning
JOURNAL      J. Eukaryot. Microbiol. 47 (1), 1-6 (2000)
MEDLINE      20114709
PUBMED       10651287
REFERENCE    2
AUTHORS      Dessen,P., Zagulski,M., Gromadka,R., Plattner,H., Kismehl,R.,
              Meyer,E., Betermier,M., Schultz,J.E., Linder,J.U., Pearlman,R.E.,
              Kung,C., Forney,J., Satir,B.H., Van Houten,J.L., Keller,A.M.,
              Froissard,M., Sperling,L. and Cohen,J.
TITLE        Parametium genome survey: a pilot project
JOURNAL      Trends Genet. 17 (6), 306-308 (2001)
MEDLINE      21273563
PUBMED       11377780
REFERENCE    3
AUTHORS      Gromadka,R. and Zagulski,M.
TITLE        Random sequencing of the Parametium macronuclear genome
JOURNAL      Unpublished
REMARK       Institute of Biochemistry and Biophysics, Polish Academy of
              Sciences, Warsaw, Poland
REFERENCE    4
AUTHORS      Cohen,J. and Sperling,L.
TITLE        Direct Submission
JOURNAL      Submitted (01-NOV-2000) Parametium Genome Survey Project, Centre de
              Genetique Moleculaire, Centre National de la Recherche
              Scientifique, 91198 Gif-sur-Yvette, France. E-mail:
              sperling@cgm.cnrs-gif.fr
COMMENT      The present survey of the Parametium tetraurelia macronuclear
              genome consists of end sequences of a library of random 4-12 kb
              fragments obtained by Sau3A partial digestion of macronuclear DNA
              cloned in the BamHI site of pSILKS-. See [4].
              Genes are predicted from matches to other sequences. For more
              information about this sequence or the Parametium Project, see
              http://parametium.cgm.cnrs-gif.fr.
FEATURES     source
              1..483
              Location/Qualifiers
                /organism="Parametium tetraurelia"
                /macronuclear
                /mol_type="genomic DNA"
                /strain="stock d4-2"
                /db_xref="taxon:5888"
ORIGIN
Query Match      10.7%; Score 18; DB 9; Length 483;
Best Local Similarity 100.0%; Pred.No. 69;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      137 AATTTCGAGGATGAGA 154
        |||||||
Db      418 AATTTCGAGGATGAGA 401

```

Search completed: August 26, 2005, 23:31:23
 Job time : 2015.75 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 26, 2005, 17:02:31 ; Search time 117.487 Seconds
(without alignments)
2465.133 Million cell updates/sec

Title: us-10-617-978-14_COPY_64_240

Perfect score: 177

Sequence: 1 gctgacgtccggaaacta.....atggaacgtgaaggtctga 177

Scoring table: OLIGO NUC

Gapop 60.0 , Gapext 60.0

Searched: 1202784 seqs, 818138359 residues

Word size : 0

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued Patents NA.*

- 1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
- 2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
- 3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
- 4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
- 5: /cgn2_6/ptodata/1/ina/PTUS_COMB.seq.*
- 6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	17	9.6	396	4	US-09-640-173-93
C 2	17	9.6	396	4	US-09-713-550-93
C 3	17	9.6	396	4	US-09-825-294-93
C 4	17	9.6	396	4	US-09-970-966-93
C 5	17	9.6	1083	4	US-09-602-777A-259
C 6	17	9.6	3662	4	US-09-907-794A-289
C 7	17	9.6	3662	4	US-09-905-125A-289
C 8	17	9.6	3662	4	US-09-902-775A-289
C 9	17	9.6	3662	4	US-09-906-700-289
C 10	17	9.6	3662	4	US-09-903-603A-289
C 11	17	9.6	3662	4	US-09-904-920A-289
C 12	17	9.6	3662	4	US-09-909-084-289
C 13	17	9.6	3662	4	US-09-905-381A-289
C 14	17	9.6	3662	4	US-09-906-618-289
C 15	17	9.6	4053	4	US-09-907-794A-293
C 16	17	9.6	4053	4	US-09-905-125A-293
C 17	17	9.6	4053	4	US-09-902-775A-293
C 18	17	9.6	4053	4	US-09-906-700-293
C 19	17	9.6	4053	4	US-09-903-603A-293
C 20	17	9.6	4053	4	US-09-904-920A-293
C 21	17	9.6	4053	4	US-09-909-064-293
C 22	17	9.6	4053	4	US-09-905-381A-293
C 23	17	9.6	4053	4	US-09-906-618-293
C 24	17	9.6	145241	4	US-09-949-016-17394
C 25	17	9.6	145241	4	US-09-949-016-17395
C 26	16	9.0	601	4	US-09-949-016-157499
C 27	16	9.0	601	4	US-09-949-016-157606

Sequence 2697, Ap
Sequence 6, Appli
Sequence 17042, A
Sequence 17043, A
Sequence 12535, A
Sequence 17476, A
Sequence 19, Appl
Sequence 19, Appl
Sequence 15, Appl
Sequence 15, Appl
Sequence 11969, A
Sequence 11930, A
Sequence 16601, A
Sequence 2, Appli
Sequence 17565, A
Sequence 16137, A
Sequence 16138, A
Sequence 16664, A

28 16 9.0 804 4 US-09-107-532A-2697
C 29 16 9.0 3655 3 US-08-878-474-6
C 30 16 9.0 12309 4 US-09-949-016-17042
C 31 16 9.0 12309 4 US-09-949-016-17043
C 32 16 9.0 29067 4 US-09-949-016-12535
C 33 16 9.0 29093 4 US-09-949-016-17476
C 34 16 9.0 35100 1 PCT-US93-06251-19
C 35 16 9.0 35100 5 PCT-US93-06251-19
C 36 16 9.0 40352 3 US-08-846-111D-15
C 37 16 9.0 40352 3 US-09-443-077-15
C 38 16 9.0 47787 4 US-09-949-016-11969
C 39 16 9.0 86947 4 US-09-949-016-11930
C 40 16 9.0 90724 4 US-09-949-016-16601
C 41 16 9.0 99960 4 US-09-762-311-2
C 42 16 9.0 116955 4 US-09-949-016-17565
C 43 16 9.0 126176 4 US-09-949-016-16137
C 44 16 9.0 126176 4 US-09-949-016-16138
C 45 16 9.0 130298 4 US-09-949-016-16664

ALIGNMENTS

RESULT 1

US-09-640-173-93/c
; Sequence 93, Application US/09640173
; Patent No. 6613515

GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun

; APPLICANT: Stolk, John A.
; TITLE OF INVENTION: OVARIAN TUMOR SEQUENCES AND

; TITLE OF INVENTION: METHODS OF USE THEREFOR
; FILE REFERENCE: 210121.484C2

; CURRENT APPLICATION NUMBER: US/09/640,173
; CURRENT FILING DATE: 2000-08-15

; NUMBER OF SEQ ID NOS: 196
; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 93
; LENGTH: 396

; TYPE: DNA

; ORGANISM: Homo sapien
; FEATURE:

; NAME/KEY: misc feature
; LOCATION: (1)..(396)

; OTHER INFORMATION: n = A,T,C or G
US-09-640-173-93

Query Match 9.6%; Score 17; DB 4; Length 396;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 149 TTCTGAAGGATGAGAAC 165
|||||
Db 199 TTCTGAAGGATGAGAAC 183

RESULT 2

US-09-713-550-93/c
; Sequence 93, Application US/09713550
; Patent No. 6617109

GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun

; APPLICANT: Stolk, John A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE

; TITLE OF INVENTION: THERAPY AND DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210121.484C4

; CURRENT APPLICATION NUMBER: US/09/713,550
; CURRENT FILING DATE: 2000-11-14

; NUMBER OF SEQ ID NOS: 205
; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 93
; LENGTH: 396

; TYPE: DNA

```
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(396)
; OTHER INFORMATION: n = A,T,C or G
US-09-713-550-93

Query Match          9.6%; Score 17; DB 4; Length 396;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 149 TTCTGAAGGATGAGAAC 165
Db 199 TTCTGAAGGATGAGAAC 183

RESULT 3
US-09-825-294-93/c
; Sequence 93, Application US/09825294
; Patent No. 6710170
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Stolk, John A.
; APPLICANT: Algate, Paul A.
; APPLICANT: Fling, Steven P.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; TREATMENT OF OVARIAN CANCER
; FILE REFERENCE: 210121.484C5
; CURRENT APPLICATION NUMBER: US/09/825,294
; CURRENT FILING DATE: 2001-04-03
; NUMBER OF SEQ ID NOS: 215
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 93
; LENGTH: 396
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(396)
; OTHER INFORMATION: n = A,T,C or G
US-09-825-294-93

Query Match          9.6%; Score 17; DB 4; Length 396;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 149 TTCTGAAGGATGAGAAC 165
Db 199 TTCTGAAGGATGAGAAC 183

RESULT 4
US-09-970-966-93/c
; Sequence 93, Application US/09970966
; Patent No. 6720146
; GENERAL INFORMATION:
; APPLICANT: Stolk, John A.
; APPLICANT: Molesh, David Alan
; APPLICANT: Fling, Steven P.
; APPLICANT: Xu, Jiangchun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; OF OVARIAN CANCER
; FILE REFERENCE: 210121.484C6
; CURRENT APPLICATION NUMBER: US/09/970,966
; CURRENT FILING DATE: 2001-10-02
; NUMBER OF SEQ ID NOS: 215
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 93
; LENGTH: 396
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(396)
; OTHER INFORMATION: n = A,T,C or G
US-09-970-966-93

Query Match          9.6%; Score 17; DB 4; Length 396;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 149 TTCTGAAGGATGAGAAC 165
Db 199 TTCTGAAGGATGAGAAC 183

RESULT 5
US-09-602-777A-259/c
; Sequence 259, Application US/09602777A
; Patent No. 6831165
; GENERAL INFORMATION:
; APPLICANT: Pompeius, Markus
; APPLICANT: Kroger, Burkhard
; APPLICANT: Schroder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Haberhauer, Gregor
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS
; INVOLVED IN HOMEOSTASIS AND ADAPTATION
; FILE REFERENCE: BGI-128CP
; CURRENT APPLICATION NUMBER: US/09/602,777A
; CURRENT FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 60/141031
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: DE 19931636.8
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19932125.6
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932126.4
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932127.2
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932128.0
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932129.9
; PRIOR FILING DATE: 1999-07-19
; PRIOR APPLICATION NUMBER: DE 19932226.0
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932920.6
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19932922.2
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19932924.9
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19932928.1
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19932930.3
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19932933.8
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19932935.4
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19932973.7
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19933002.6
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19933003.4
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19933005.0
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19933006.9
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19941378.9
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: DE 19941379.7
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: DE 19941390.8
```

; PRIOR FILING DATE: 1999-08-31
 ; PRIOR APPLICATION NUMBER: DE 19941391.6
 ; PRIOR FILING DATE: 1999-08-31
 ; PRIOR APPLICATION NUMBER: DE 19942088.2
 ; PRIOR FILING DATE: 1999-09-03
 ; NUMBER OF SEQ ID NOS: 442
 ; SEQ ID NO 259
 ; LENGTH: 1083
 ; TYPE: DNA
 ; ORGANISM: Corynebacterium glutamicum
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (101)..(1060)
 ; OTHER INFORMATION: RXN01466
 US-09-602-777A-259

Query Match 9.6%; Score 17; DB 4; Length 1083;
 Best Local Similarity 100.0%; Pred. No. 12;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 127 GCCTTCCAATGCTGGTG 143
 |||||
 Db 877 GCCTTCCAATGCTGGTG 861

RESULT 6
 US-09-907-794A-289/c
 ; Sequence 289, Application US/09907794A
 ; Patent No. 6635468
 ; GENERAL INFORMATION:
 ; APPLICANT: Genentech, Inc.
 ; APPLICANT: Ashkenazi, Avi
 ; APPLICANT: Botstein, David
 ; APPLICANT: Desnovers, Luc
 ; APPLICANT: Eaton, Dan L.
 ; APPLICANT: Ferrara, Napoleone
 ; APPLICANT: Filvaroff, Ellen
 ; APPLICANT: Fong, Sherman
 ; APPLICANT: Gao, Wei-Qiang
 ; APPLICANT: Gerber, Hanspeter
 ; APPLICANT: Gerritsen, Mary E.
 ; APPLICANT: Goddard, A.
 ; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Grimaldi, Christopher J.
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Hillan, Kenneth, J.
 ; APPLICANT: Kljavin, Ivar J.
 ; APPLICANT: Mather, Jennie P.
 ; APPLICANT: Pan, James
 ; APPLICANT: Paoni, Nicholas F.
 ; APPLICANT: Roy, Margaret Ann
 ; APPLICANT: Stewart, Timothy A.
 ; APPLICANT: Tumas, Daniel
 ; APPLICANT: Williams, P. Mickey
 ; APPLICANT: Wood, William, I.
 ; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
 ; FILE REFERENCE: 10466-14
 ; CURRENT APPLICATION NUMBER: US/09/907,794A
 ; CURRENT FILING DATE: 2001-07-17
 ; PRIOR APPLICATION NUMBER: PCT/US00/04414
 ; PRIOR FILING DATE: 2000-02-22
 ; PRIOR APPLICATION NUMBER: US 60/143,048
 ; PRIOR FILING DATE: 1999-07-07
 ; PRIOR APPLICATION NUMBER: US 60/145,698
 ; PRIOR FILING DATE: 1999-07-26
 ; PRIOR APPLICATION NUMBER: US 60/146,222
 ; PRIOR FILING DATE: 1999-07-28
 ; PRIOR APPLICATION NUMBER: PCT/US99/20594
 ; PRIOR FILING DATE: 1999-09-08
 ; PRIOR APPLICATION NUMBER: PCT/US99/20944
 ; PRIOR FILING DATE: 1999-09-13
 ; PRIOR APPLICATION NUMBER: PCT/US99/21090

; PRIOR FILING DATE: 1999-09-15
 ; PRIOR APPLICATION NUMBER: PCT/US99/21547
 ; PRIOR FILING DATE: 1999-09-15
 ; PRIOR APPLICATION NUMBER: PCT/US99/23089
 ; PRIOR FILING DATE: 1999-10-05
 ; PRIOR APPLICATION NUMBER: PCT/US99/28214
 ; PRIOR FILING DATE: 1999-11-29
 ; PRIOR APPLICATION NUMBER: PCT/US99/28313
 ; PRIOR FILING DATE: 1999-11-30
 ; PRIOR APPLICATION NUMBER: PCT/US99/28564
 ; PRIOR FILING DATE: 1999-12-02
 ; PRIOR APPLICATION NUMBER: PCT/US99/28565
 ; PRIOR FILING DATE: 1999-12-02
 ; PRIOR APPLICATION NUMBER: PCT/US99/30095
 ; PRIOR FILING DATE: 1999-12-16
 ; PRIOR APPLICATION NUMBER: PCT/US99/30911
 ; PRIOR FILING DATE: 1999-12-20
 ; PRIOR APPLICATION NUMBER: PCT/US99/30999
 ; PRIOR FILING DATE: 1999-12-20
 ; PRIOR APPLICATION NUMBER: PCT/US00/00219
 ; PRIOR FILING DATE: 2000-01-05
 ; NUMBER OF SEQ ID NOS: 423
 ; SEQ ID NO 289
 ; LENGTH: 3662
 ; TYPE: DNA
 ; ORGANISM: Homo Sapien
 US-09-907-794A-289

Query Match 9.6%; Score 17; DB 4; Length 3662;
 Best Local Similarity 100.0%; Pred. No. 12;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 149 TTCTGAAGGATGAGAAC 165
 |||||
 Db 2818 TTCTGAAGGATGAGAAC 2802

RESULT 7
 US-09-905-125A-289/c
 ; Sequence 289, Application US/09905125A
 ; Patent No. 6664376
 ; GENERAL INFORMATION:
 ; APPLICANT: Genentech, Inc.
 ; APPLICANT: Ashkenazi, Avi
 ; APPLICANT: Botstein, David
 ; APPLICANT: Desnovers, Luc
 ; APPLICANT: Eaton, Dan L.
 ; APPLICANT: Ferrara, Napoleone
 ; APPLICANT: Filvaroff, Ellen
 ; APPLICANT: Fong, Sherman
 ; APPLICANT: Gao, Wei-Qiang
 ; APPLICANT: Gerber, Hanspeter
 ; APPLICANT: Gerritsen, Mary E.
 ; APPLICANT: Goddard, A.
 ; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Grimaldi, Christopher J.
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Hillan, Kenneth, J.
 ; APPLICANT: Kljavin, Ivar J.
 ; APPLICANT: Mather, Jennie P.
 ; APPLICANT: Pan, James
 ; APPLICANT: Paoni, Nicholas F.
 ; APPLICANT: Roy, Margaret Ann
 ; APPLICANT: Stewart, Timothy A.
 ; APPLICANT: Tumas, Daniel
 ; APPLICANT: Williams, P. Mickey
 ; APPLICANT: Wood, William, I.
 ; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
 ; FILE REFERENCE: 10466-14
 ; CURRENT APPLICATION NUMBER: US/09/905,125A
 ; CURRENT FILING DATE: 2001-07-12
 ; PRIOR APPLICATION NUMBER: PCT/US00/04414

; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 289
; LENGTH: 3662
; TYPE: DNA
; ORGANISM: Homo Sapien
US-09-905-125A-289

Query Match 9.6%; Score 17; DB 4; Length 3662;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 149 TTCTGAAGGATGAGAAC 165
Db 2818 TTCTGAAGGATGAGAAC 2802

RESULT 8
US-09-902-775A-289/c
; Sequence 289, Application US/09902775A
; Patent No. 6686451
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James

; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/902,775A
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 289
; LENGTH: 3662
; TYPE: DNA
; ORGANISM: Homo Sapien
US-09-902-775A-289

Query Match 9.6%; Score 17; DB 4; Length 3662;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 149 TTCTGAAGGATGAGAAC 165
Db 2818 TTCTGAAGGATGAGAAC 2802

RESULT 9
US-09-906-700-289/c
; Sequence 289, Application US/09906700
; Patent No. 6723535
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen

APPLICANT: Fong, Sherman
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerber, Hanspeter
APPLICANT: Geritsen, Mary E.
APPLICANT: Goddard, A.
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, Christopher J.
APPLICANT: Gurney, Austin L.
APPLICANT: Hillan, Kenneth, J.
APPLICANT: Kljavin, Ivar J.
APPLICANT: Mather, Jennie P.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William, I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: 10466-14
CURRENT APPLICATION NUMBER: US/09/906,700
PRIOR FILING DATE: 2000-09-18
PRIOR APPLICATION NUMBER: PCT/US00/04414
PRIOR FILING DATE: 2000-02-22
PRIOR APPLICATION NUMBER: US 60/143,048
PRIOR FILING DATE: 1999-07-07
PRIOR APPLICATION NUMBER: US 60/145,698
PRIOR FILING DATE: 1999-07-26
PRIOR APPLICATION NUMBER: US 60/146,222
PRIOR FILING DATE: 1999-07-28
PRIOR APPLICATION NUMBER: PCT/US99/20594
PRIOR FILING DATE: 1999-09-08
PRIOR APPLICATION NUMBER: PCT/US99/20944
PRIOR FILING DATE: 1999-09-13
PRIOR APPLICATION NUMBER: PCT/US99/21090
PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: PCT/US99/21547
PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: PCT/US99/23089
PRIOR FILING DATE: 1999-10-05
PRIOR APPLICATION NUMBER: PCT/US99/28214
PRIOR FILING DATE: 1999-11-29
PRIOR APPLICATION NUMBER: PCT/US99/28313
PRIOR FILING DATE: 1999-11-30
PRIOR APPLICATION NUMBER: PCT/US99/28564
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US99/28565
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: PCT/US99/30095
PRIOR FILING DATE: 1999-12-20
PRIOR APPLICATION NUMBER: PCT/US99/30999
PRIOR FILING DATE: 1999-12-20
PRIOR APPLICATION NUMBER: PCT/US00/00219
PRIOR FILING DATE: 2000-01-05
NUMBER OF SEQ ID NOS: 423
SEQ ID NO 289
LENGTH: 3662
TYPE: DNA
ORGANISM: Homo Sapien

US-09-906-700-289

Query Match 9.6%; Score 17; DB 4; Length 3662;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 149 TTCTGAAGATGAGAAC 165
|||||
Db 2818 TTCTGAAGATGAGAAC 2802

RESULT 10
US-09-903-603A-289/c
Sequence 289, Application US/09903603A
Patent No. 6767995
GENERAL INFORMATION:
APPLICANT: Genentech, Inc.
APPLICANT: Ashkenazi, Avi
APPLICANT: Botstein, David
APPLICANT: Deanoyers, Luc
APPLICANT: Eaton, Dan L.
APPLICANT: Ferrara, Napoleone
APPLICANT: Filvaroff, Ellen
APPLICANT: Fong, Sherman
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerber, Hanspeter
APPLICANT: Geritsen, Mary E.
APPLICANT: Goddard, A.
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, Christopher J.
APPLICANT: Gurney, Austin L.
APPLICANT: Hillan, Kenneth, J.
APPLICANT: Kljavin, Ivar J.
APPLICANT: Mather, Jennie P.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William, I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: GNE.16182C12
CURRENT APPLICATION NUMBER: US/09/903,603A
CURRENT FILING DATE: 2001-07-11
PRIOR APPLICATION NUMBER: PCT/US00/04414
PRIOR FILING DATE: 2000-02-22
PRIOR APPLICATION NUMBER: US 60/143,048
PRIOR FILING DATE: 1999-07-07
PRIOR APPLICATION NUMBER: US 60/145,698
PRIOR FILING DATE: 1999-07-26
PRIOR APPLICATION NUMBER: US 60/146,222
PRIOR FILING DATE: 1999-07-28
PRIOR APPLICATION NUMBER: PCT/US99/20594
PRIOR FILING DATE: 1999-09-08
PRIOR APPLICATION NUMBER: PCT/US99/20944
PRIOR FILING DATE: 1999-09-13
PRIOR APPLICATION NUMBER: PCT/US99/21090
PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: PCT/US99/21547
PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: PCT/US99/23089
PRIOR FILING DATE: 1999-10-05
PRIOR APPLICATION NUMBER: PCT/US99/28214
PRIOR FILING DATE: 1999-11-29
PRIOR APPLICATION NUMBER: PCT/US99/28313
PRIOR FILING DATE: 1999-11-30
PRIOR APPLICATION NUMBER: PCT/US99/28564
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US99/28565
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US99/30095
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: PCT/US99/30911
PRIOR FILING DATE: 1999-12-20
PRIOR APPLICATION NUMBER: PCT/US99/30999
PRIOR FILING DATE: 1999-12-20
PRIOR APPLICATION NUMBER: PCT/US00/00219
PRIOR FILING DATE: 2000-01-05
NUMBER OF SEQ ID NOS: 423
SEQ ID NO 289
LENGTH: 3662
TYPE: DNA

; ORGANISM: Homo Sapien
US-09-903-603A-289

Query Match 9.6%; Score 17; DB 4; Length 3662;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 149 TTCTGAAGGATGAGAAC 165
|||||
Db 2818 TTCTGAAGGATGAGAAC 2802

RESULT 11

US-09-904-920A-289/c
; Sequence 289, Application US/09904920A

; Patent No. 6806352

; GENERAL INFORMATION:

; APPLICANT: Genentech, Inc.

; APPLICANT: Ashkenazi, Avi

; APPLICANT: Botstein, David

; APPLICANT: Desnovers, Luc

; APPLICANT: Eaton, Dan L.

; APPLICANT: Ferrara, Napoleone

; APPLICANT: Filvaroff, Ellen

; APPLICANT: Fong, Sherman

; APPLICANT: Gao, Wei-Qiang

; APPLICANT: Gerber, Hanspeter

; APPLICANT: Gerriksen, Mary E.

; APPLICANT: Goddard, A.

; APPLICANT: Godowski, Paul J.

; APPLICANT: Grimaldi, Christopher J.

; APPLICANT: Gurney, Austin L.

; APPLICANT: Hillan, Kenneth, J.

; APPLICANT: Kljavin, Ivar J.

; APPLICANT: Mather, Jennie P.

; APPLICANT: Pan, James

; APPLICANT: Páoni, Nicholas F.

; APPLICANT: Róy, Margaret Ann

; APPLICANT: Stewart, Timothy A.

; APPLICANT: Tumas, Daniel

; APPLICANT: Williams, P. Mickey

; APPLICANT: Wood, William, I.

; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic

; FILE REFERENCE: 10466-14

; CURRENT APPLICATION NUMBER: US/09/904,920A

; PRIOR FILING DATE: 2001-07-13

; PRIOR APPLICATION NUMBER: PCT/US00/04414

; PRIOR FILING DATE: 2000-02-22

; PRIOR APPLICATION NUMBER: US 60/143,048

; PRIOR FILING DATE: 1999-07-07

; PRIOR APPLICATION NUMBER: US 60/145,698

; PRIOR FILING DATE: 1999-07-26

; PRIOR APPLICATION NUMBER: US 60/146,222

; PRIOR FILING DATE: 1999-07-28

; PRIOR APPLICATION NUMBER: PCT/US99/20594

; PRIOR FILING DATE: 1999-09-08

; PRIOR APPLICATION NUMBER: PCT/US99/20944

; PRIOR FILING DATE: 1999-09-13

; PRIOR APPLICATION NUMBER: PCT/US99/21090

; PRIOR FILING DATE: 1999-09-15

; PRIOR APPLICATION NUMBER: PCT/US99/21547

; PRIOR FILING DATE: 1999-09-15

; PRIOR APPLICATION NUMBER: PCT/US99/23089

; PRIOR FILING DATE: 1999-10-05

; PRIOR APPLICATION NUMBER: PCT/US99/28214

; PRIOR FILING DATE: 1999-11-29

; PRIOR APPLICATION NUMBER: PCT/US99/28313

; PRIOR FILING DATE: 1999-11-30

; PRIOR APPLICATION NUMBER: PCT/US99/28564

; PRIOR FILING DATE: 1999-12-02

; PRIOR APPLICATION NUMBER: PCT/US99/28565

; PRIOR FILING DATE: 1999-12-02

; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 289
; LENGTH: 3662
; TYPE: DNA
; ORGANISM: Homo Sapien
US-09-904-920A-289

Query Match 9.6%; Score 17; DB 4; Length 3662;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 149 TTCTGAAGGATGAGAAC 165
|||||
Db 2818 TTCTGAAGGATGAGAAC 2802

RESULT 12

US-09-909-064-289/c
; Sequence 289, Application US/09909064

; Patent No. 6818449

; GENERAL INFORMATION:

; APPLICANT: Genentech, Inc.

; APPLICANT: Ashkenazi, Avi

; APPLICANT: Botstein, David

; APPLICANT: Desnovers, Luc

; APPLICANT: Eaton, Dan L.

; APPLICANT: Ferrara, Napoleone

; APPLICANT: Filvaroff, Ellen

; APPLICANT: Fong, Sherman

; APPLICANT: Gao, Wei-Qiang

; APPLICANT: Gerber, Hanspeter

; APPLICANT: Gerriksen, Mary E.

; APPLICANT: Goddard, A.

; APPLICANT: Godowski, Paul J.

; APPLICANT: Grimaldi, Christopher J.

; APPLICANT: Gurney, Austin L.

; APPLICANT: Hillan, Kenneth, J.

; APPLICANT: Kljavin, Ivar J.

; APPLICANT: Mather, Jennie P.

; APPLICANT: Pan, James

; APPLICANT: Paoni, Nicholas F.

; APPLICANT: Roy, Margaret Ann

; APPLICANT: Stewart, Timothy A.

; APPLICANT: Tumas, Daniel

; APPLICANT: Williams, P. Mickey

; APPLICANT: Wood, William, I.

; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic

; FILE REFERENCE: 10466-14

; CURRENT APPLICATION NUMBER: US/09/909,064

; PRIOR FILING DATE: 2001-07-18

; PRIOR APPLICATION NUMBER: PCT/US00/04414

; PRIOR FILING DATE: 2000-02-22

; PRIOR APPLICATION NUMBER: US 60/143,048

; PRIOR FILING DATE: 1999-07-07

; PRIOR APPLICATION NUMBER: US 60/145,698

; PRIOR FILING DATE: 1999-07-26

; PRIOR APPLICATION NUMBER: US 60/146,222

; PRIOR FILING DATE: 1999-07-28

; PRIOR APPLICATION NUMBER: PCT/US99/20594

; PRIOR FILING DATE: 1999-09-08

; PRIOR APPLICATION NUMBER: PCT/US99/20944

; PRIOR FILING DATE: 1999-09-13

; PRIOR APPLICATION NUMBER: PCT/US99/21090

; PRIOR FILING DATE: 1999-09-15

;; PRIOR APPLICATION NUMBER: PCT/US99/21547
;; PRIOR FILING DATE: 1999-09-15
;; PRIOR APPLICATION NUMBER: PCT/US99/23089
;; PRIOR FILING DATE: 1999-10-05
;; PRIOR APPLICATION NUMBER: PCT/US99/28214
;; PRIOR FILING DATE: 1999-11-29
;; PRIOR APPLICATION NUMBER: PCT/US99/28313
;; PRIOR FILING DATE: 1999-11-30
;; PRIOR APPLICATION NUMBER: PCT/US99/28564
;; PRIOR FILING DATE: 1999-12-02
;; PRIOR APPLICATION NUMBER: PCT/US99/28565
;; PRIOR FILING DATE: 1999-12-02
;; PRIOR APPLICATION NUMBER: PCT/US99/30095
;; PRIOR FILING DATE: 1999-12-16
;; PRIOR APPLICATION NUMBER: PCT/US99/30911
;; PRIOR FILING DATE: 1999-12-20
;; PRIOR APPLICATION NUMBER: PCT/US99/30999
;; PRIOR FILING DATE: 1999-12-20
;; PRIOR APPLICATION NUMBER: PCT/US00/00219
;; PRIOR FILING DATE: 2000-01-05
;; NUMBER OF SEQ ID NOS: 423
;; SEQ ID NO 289
;; LENGTH: 3662
;; TYPE: DNA
;; ORGANISM: Homo Sapien
US-09-909-064-289

Query Match 9.6%; Score 17; DB 4; Length 3662;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 149 TTCTGAAGGATGAGAAC 165
|||||
Db 2818 TTCTGAAGGATGAGAAC 2802

RESULT 13
US-09-905-381A-289/c
; Sequence 289, Application US/09905381A
; Patent No. 6818746
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnovers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Pao, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/905,381A
; CURRENT FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22

;; PRIOR APPLICATION NUMBER: US 60/143,048
;; PRIOR FILING DATE: 1999-07-07
;; PRIOR APPLICATION NUMBER: US 60/145,698
;; PRIOR FILING DATE: 1999-07-26
;; PRIOR APPLICATION NUMBER: US 60/146,222
;; PRIOR FILING DATE: 1999-07-28
;; PRIOR APPLICATION NUMBER: PCT/US99/20594
;; PRIOR FILING DATE: 1999-09-08
;; PRIOR APPLICATION NUMBER: PCT/US99/20944
;; PRIOR FILING DATE: 1999-09-13
;; PRIOR APPLICATION NUMBER: PCT/US99/21090
;; PRIOR FILING DATE: 1999-09-15
;; PRIOR APPLICATION NUMBER: PCT/US99/21547
;; PRIOR FILING DATE: 1999-09-15
;; PRIOR APPLICATION NUMBER: PCT/US99/23089
;; PRIOR FILING DATE: 1999-10-05
;; PRIOR APPLICATION NUMBER: PCT/US99/28214
;; PRIOR FILING DATE: 1999-11-29
;; PRIOR APPLICATION NUMBER: PCT/US99/28313
;; PRIOR FILING DATE: 1999-11-30
;; PRIOR APPLICATION NUMBER: PCT/US99/28564
;; PRIOR FILING DATE: 1999-12-02
;; PRIOR APPLICATION NUMBER: PCT/US99/28565
;; PRIOR FILING DATE: 1999-12-02
;; PRIOR APPLICATION NUMBER: PCT/US99/30095
;; PRIOR FILING DATE: 1999-12-16
;; PRIOR APPLICATION NUMBER: PCT/US99/30911
;; PRIOR FILING DATE: 1999-12-20
;; PRIOR APPLICATION NUMBER: PCT/US99/30999
;; PRIOR FILING DATE: 1999-12-20
;; PRIOR APPLICATION NUMBER: PCT/US00/00219
;; PRIOR FILING DATE: 2000-01-05
;; NUMBER OF SEQ ID NOS: 423
;; SEQ ID NO 289
;; LENGTH: 3662
;; TYPE: DNA
;; ORGANISM: Homo Sapien
US-09-905-381A-289

Query Match 9.6%; Score 17; DB 4; Length 3662;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 149 TTCTGAAGGATGAGAAC 165
|||||
Db 2818 TTCTGAAGGATGAGAAC 2802

RESULT 14
US-09-906-618-289/c
; Sequence 289, Application US/09906618
; Patent No. 6828146
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnovers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Pao, Nicholas F.

```
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/906,618
; CURRENT FILING DATE: 2001-07-16
; PRIOR FILING DATE: 2000-02-22
; PRIOR FILING DATE: 2000-02-22
; PRIOR FILING DATE: 1999-07-07
; PRIOR FILING DATE: 1999-07-26
; PRIOR FILING DATE: 1999-07-26
; PRIOR FILING DATE: 1999-07-28
; PRIOR FILING DATE: 1999-07-28
; PRIOR FILING DATE: 1999-09-08
; PRIOR FILING DATE: 1999-09-13
; PRIOR FILING DATE: 1999-09-15
; PRIOR FILING DATE: 1999-09-15
; PRIOR FILING DATE: 1999-09-15
; PRIOR FILING DATE: 1999-10-05
; PRIOR FILING DATE: 1999-11-29
; PRIOR FILING DATE: 1999-11-29
; PRIOR FILING DATE: 1999-11-30
; PRIOR FILING DATE: 1999-11-30
; PRIOR FILING DATE: 1999-12-02
; PRIOR FILING DATE: 1999-12-02
; PRIOR FILING DATE: 1999-12-02
; PRIOR FILING DATE: 1999-12-16
; PRIOR FILING DATE: 1999-12-20
; PRIOR FILING DATE: 1999-12-20
; PRIOR FILING DATE: 1999-12-20
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 289
; LENGTH: 3662
; TYPE: DNA
; ORGANISM: Homo Sapien
US-09-906-618-289
```

```
Query Match          9.6%; Score 17; DB 4; Length 3662;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 149 TTCTGAAGGATGAGAAC 165
||| ||||| ||||| |||||
Db 2818 TTCTGAAGGATGAGAAC 2802
```

```
RESULT 15
US-09-907-794A-293/c
; Sequence 293, Application US/09907794A
; Patent No. 6635468
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnovers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
```

```
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kijavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/907,794A
; CURRENT FILING DATE: 2001-07-17
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 293
; LENGTH: 4053
; TYPE: DNA
; ORGANISM: Homo Sapien
US-09-907-794A-293
```

```
Query Match          9.6%; Score 17; DB 4; Length 4053;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 149 TTCTGAAGGATGAGAAC 165
||| ||||| ||||| |||||
Db 3180 TTCTGAAGGATGAGAAC 3164
```

Search completed: August 26, 2005, 18:46:52

Job time : 118.487 secs



This Page Blank (uspio)